

PATENT COOPERATION TREATY

PCT

NOTIFICATION OF ELECTION

(PCT Rule 61.2)

From the INTERNATIONAL BUREAU

To:

Assistant Commissioner for Patents
United States Patent and Trademark
Office
Box PCT
Washington, D.C.20231
ETATS-UNIS D'AMERIQUE

in its capacity as elected Office

Date of mailing (day/month/year) 14 September 2000 (14.09.00)	
International application No. PCT/JP00/00386	Applicant's or agent's file reference SK00PCT9
International filing date (day/month/year) 26 January 2000 (26.01.00)	Priority date (day/month/year) 26 January 1999 (26.01.99)
Applicant TAKAHASHI, Yasushi	

1. The designated Office is hereby notified of its election made:

☒ in the demand filed with the International Preliminary Examining Authority on:
25 July 2000 (25.07.00)

☐ in a notice effecting later election filed with the International Bureau on:

2. The election ☒ was
☐ was not

made before the expiration of 19 months from the priority date or, where Rule 32 applies, within the time limit under Rule 32.2(b).

BEST AVAILABLE COPY

The International Bureau of WIPO 34, chemin des Colombettes 1211 Geneva 20, Switzerland	Authorized officer Christelle Croci
Facsimile No.: (41-22) 740.14.35	Telephone No.: (41-22) 338.83.38

PCT

INTERNATIONAL SEARCH REPORT

(PCT Article 18 and Rules 43 and 44)

Applicant's or agent's file reference P/23222WO/MW	FOR FURTHER ACTION see Notification of Transmittal of International Search Report (Form PCT/ISA/220) as well as, where applicable, item 5 below.	
International application No. PCT/IB 00/ 00135	International filing date (day/month/year) 28/01/2000	(Earliest) Priority Date (day/month/year) 01/02/1999
Applicant EMUSE CORPORATION et al.		

This International Search Report has been prepared by this International Searching Authority and is transmitted to the applicant according to Article 18. A copy is being transmitted to the International Bureau.

This International Search Report consists of a total of 6 sheets.



It is also accompanied by a copy of each prior art document cited in this report.

1. Basis of the report

- a. With regard to the **language**, the international search was carried out on the basis of the international application in the language in which it was filed, unless otherwise indicated under this item.



the international search was carried out on the basis of a translation of the international application furnished to this Authority (Rule 23.1(b)).

- b. With regard to any **nucleotide and/or amino acid sequence** disclosed in the international application, the international search was carried out on the basis of the sequence listing :



contained in the international application in written form.



filed together with the international application in computer readable form.



furnished subsequently to this Authority in written form.



furnished subsequently to this Authority in computer readable form.



the statement that the subsequently furnished written sequence listing does not go beyond the disclosure in the international application as filed has been furnished.



the statement that the information recorded in computer readable form is identical to the written sequence listing has been furnished

2.



Certain claims were found unsearchable (See Box I).

3.



Unity of invention is lacking (see Box II).

4. With regard to the **title**,

the text is approved as submitted by the applicant.



the text has been established by this Authority to read as follows:

5. With regard to the **abstract**,

the text is approved as submitted by the applicant.



the text has been established, according to Rule 38.2(b), by this Authority as it appears in Box III. The applicant may, within one month from the date of mailing of this international search report, submit comments to this Authority.

6. The figure of the **drawings** to be published with the abstract is Figure No.

as suggested by the applicant.



because the applicant failed to suggest a figure.



because this figure better characterizes the invention.

1

None of the figures.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/IB 00/00135

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
2. ☐ Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:
3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. ☒ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.
- ☒ No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. Claims: 1-7,10

An interactive system including a means for providing a video programme signal, a means for generating interactive content data associated with at least one object, said data being associated with frames of said video programme signal in which the object appears, a means for multiplexing said data with the video program signal, a means for viewing the video programme signal, a means for retrieving said data and a means for using said data to obtains details of the object, said using means includes a means for producing a list of details of said object and a means for selecting from said list.

2. Claims: 8,9,27-32

An apparatus for embedding a data sequence within a generic digital transport stream, including a means for receiving a data sequence of interactive content data associated with an object in a digitised video signal, a means for synchronising the data sequence with the video and audio of the digitised video signal to generate a further transport stream, and a means for associating a packet identifier with the further transport stream, wherein the means for receiving a data sequence includes a means for receiving elementary streams comprising a digital video signal stream, a digital audio stream, a digital data sequence and a digital control data stream, a means for packetising each of the data streams into fixed sized blocks and adding a protocol header to produce packetised elementary streams, and means for synchronising the packetised elementary streams with time stamps to establish a relationship between the data streams.

3. Claims: 11-26

An apparatus for associating data representative of an object with a digital video programme including a means for providing a digital video programme having plural individual frames at least some of which incorporate said object, a means for selecting a frame of the video programme in which said object appears to provide a key-frame, a means for selecting said object within the key-frame with which data is to be associated, a means for extracting attributes of the object from the key-frame, a means for associating interactive data with the object in the key-frame, a means for utilising the attributes of the object for tracking the object through subsequent frames of the video programme, whereby said interactive data are associated with the object in subsequent frames in which the object has been tracked

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

and said interactive data are embedded with data representative of said object in a data sequence.

4. Claims: 33-37

An apparatus for retrieving data embedded in a generic digital transport stream in which the embedded data includes a data sequence of data associated with objects of the generic digital transport stream, comprising a means for recognizing a packet identifier within the transport stream, a means for extracting the data sequence from the transport stream, a means for identifying objects within the video sequence from which to retrieve associated data, a means for synchronising said data sequence to said identified objects and a means for interactively using said associated data, wherein a means is provided for selecting a frame to display the objects having embedded associated data and for selecting one of the displayed objects.

INTERNATIONAL SEARCH REPORT

International Application No

P B 00/00135

A. CLASSIFICATION OF SUBJECT MATTER

IPC 7 H04N7/24

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 H04N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

WPI Data, PAJ, INSPEC

1, 11, 27, 33

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
<p>D10</p> <p>X</p> <p>Y</p> <p>A</p>	<p>✓ WO 97 41690 A (AWARD SOFTWARE INTERNATIONAL I) 6 November 1997 (1997-11-06) page 2, line 26 -page 7, line 28 page 9, line 20 -page 16, line 18 abstract; figures 2A,2B,2C,3A,3B</p>	<p>1,2,4-6, 10</p> <p>3,7</p> <p>8,9, 11-37</p>
<p>D11</p> <p>Y</p> <p>A</p>	<p>--- ✓ EP 0 851 681 A (HITACHI LTD) 1 July 1998 (1998-07-01) column 2, line 44 -column 7, line 14 column 9, line 5 -column 19, line 54 abstract; figures 1,7A,7B,7C,10</p> <p>---</p> <p>-/--</p>	<p>3,34-36</p> <p>1,2,4,5, 8-10, 27-33</p>

☒ Further documents are listed in the continuation of box C.☒ Patent family members are listed in annex.

° Special categories of cited documents :

- *A* document defining the general state of the art which is not considered to be of particular relevance
- *E* earlier document but published on or after the international filing date
- *L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- *O* document referring to an oral disclosure, use, exhibition or other means
- *P* document published prior to the international filing date but later than the priority date claimed

- *T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- *X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- *Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- * & * document member of the same patent family

Date of the actual completion of the international search

7 August 2000

Date of mailing of the international search report

11. 09. 2000

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Authorized officer

La, V

INTERNATIONAL SEARCH REPORT

International Application No

P B 00/00135

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P12 Y	POPE A ET AL: "Video abstraction: summarizing video content for retrieval and visualization" CONFERENCE RECORD OF THIRTY-SECOND ASILOMAR CONFERENCE ON SIGNALS, SYSTEMS AND COMPUTERS (CAT. NO.98CH36284), CONFERENCE RECORD OF THIRTY-SECOND ASILOMAR CONFERENCE ON SIGNALS, SYSTEMS AND COMPUTERS, PACIFIC GROVE, CA, USA, 1-4 NOV. 1998, pages 915-919 vol.1, XP002139169 ✓ 1998, Piscataway, NJ, USA, IEEE, USA ISBN: 0-7803-5148-7 section 1."Introduction" section 2.1 "Generic schema" page 917, left-hand column, line 27 -right-hand column, line 9 page 918, left-hand column, line 6 - line 19	7,14-16, 23
A	abstract; figure 1	1,6,11
P13 X	✓ US 5 708 845 A (WISTENDAHL DOUGLASS A ET AL) 13 January 1998 (1998-01-13)	11-13, 17-19, 21,22, 24-26
Y	the whole document	14-16, 20,23
P14 Y	✓ EP 0 675 461 A (CASIO COMPUTER CO LTD) ✓ 4 October 1995 (1995-10-04) column 21, line 15 - line 29	20
A	abstract	11
P15 X	✓ WO 98 57499 A (SHARP KK) 17 December 1998 (1998-12-17)	1,8,9, 27,28, 30,32, 33,37
Y	page 6, line 2 -page 7, line 4 page 7, line 27 -page 12, line 16 abstract	29,31, 34-36
	--- -/--	

INTERNATIONAL SEARCH REPORT

International Application No

P B 00/00135

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
716 Y	SARGINSON P A: "MPEG-2: A TUTORIAL INTRODUCTION TO THE SYSTEMS LAYER" IEE COLLOQUIUM ON MPEG WHAT IT IS AND WHAT IT ISN'T, GB, IEE, LONDON, 1995, pages 4-1-4-13, XP000560804 ✓ section "The Basics" section "Packetised Elementary Streams" sections "The Transport Stream multiplex", "The transport packet header", "Programme Specific Information"	29,31
A	section "Time stamps and clock references : the basics", "Time stamps and clock references : the details"	1,8,9, 27,28, 30,32
717 X	✓ WO 98 43437 A (DECLERCK CHRISTOPHE ; CANAL PLUS SA (FR); CHAUMET CHRISTOPHE (FR)) 1 October 1998 (1998-10-01) page 1, line 1 - page 8, line 3 page 12, line 22 - page 18, line 20 abstract	1,10
A		2-5,8,9, 27-32
718 X	✓ US 5 818 935 A (MAA CHIA-YIU) 6 October 1998 (1998-10-06) column 2, line 48 - column 3, line 59 column 6, line 7 - line 39 column 14, line 61 - column 15, line 67 column 17, line 23 - column 18, line 20 column 19, line 11 - column 20, line 53 abstract	1,4,5,10
A		2,3,8,9, 27-32
719 X	--- "MULTIMEDIA HYPERVIDEO LINKS FOR FULL MOTION VIDEOS" IBM TECHNICAL DISCLOSURE BULLETIN, US, IBM CORP. NEW YORK, vol. 37, no. 4A, 1 April 1994 (1994-04-01), page 95 ✓ XP000446196 ISSN: 0018-8689 the whole document	11,18
720 A	✓ EP 0 596 823 A (IBM) 11 May 1994 (1994-05-11) the whole document	1-37
721 A	✓ WO 98 00976 A (SUN MICROSYSTEMS INC ; THOMSON CONSUMER ELECTRONICS (US)) 8 January 1998 (1998-01-08) page 3, line 15 - line 27 abstract; figures 5,6 --- -/-	1-7,10, 33-37

INTERNATIONAL SEARCH REPORT

Inter Application No

P B 00/00135

C:(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A <i>P22</i>	<p>✓ WO 97 19560 A (IMEDIA CORP) 29 May 1997 (1997-05-29) page 8, line 3 -page 11, line 29 page 14, line 3 -page 24, line 27 abstract; figures 1,4,6</p> <p>---</p>	1-10, 27-37
A <i>P23</i>	<p>OLSON C F ET AL: "AUTOMATIC TARGET RECOGNITION BY MATCHING ORIENTED EDGE PIXELS" IEEE TRANSACTIONS ON IMAGE PROCESSING,US,IEEE INC. NEW YORK, vol. 6, no. 1, 1997, pages 103-113, ✓ XP000642455 ISSN: 1057-7149 section II "Matching oriented edge pixels" abstract</p> <p>---</p>	11-26
A <i>P24</i>	<p>GUNSEL B ET AL: "TEMPORAL VIDEO SEGMENTATION USING UNSUPERVISED CLUSTERING AND SEMANTIC OBJECT TRACKING" JOURNAL OF ELECTRONIC IMAGING,US,SPIE + IS&T, vol. 7, no. 3, 1 July 1998 (1998-07-01), pages 592-604, XP000771766✓ ISSN: 1017-9909 the whole document</p> <p>---</p>	11-26
A <i>P25</i>	<p>✓ WO 98 21688 A (SARNOFF CORP) 22 May 1998 (1998-05-22) page 16, line 29 -page 20, line 10 abstract</p> <p>-----</p>	11-26

INTERNATIONAL SEARCH REPORT

Information on patent family members

Inter Application No

P B 00/00135

Patent document cited in search report		Publication date	Patent family member(s)	Publication date
WO 9741690	A	06-11-1997	US 5929849 A	27-07-1999
			CN 1221538 A	30-06-1999
			EP 0896774 A	17-02-1999
			JP 11510978 T	21-09-1999
EP 0851681	A	01-07-1998	JP 10191273 A	21-07-1998
US 5708845	A	13-01-1998	CA 2233444 A	03-04-1997
			EP 0902928 A	24-03-1999
			JP 11512902 T	02-11-1999
			WO 9712342 A	03-04-1997
EP 0675461	A	04-10-1995	JP 7262391 A	13-10-1995
			JP 7262398 A	13-10-1995
			JP 7262399 A	13-10-1995
			JP 7264472 A	13-10-1995
			US 5611037 A	11-03-1997
WO 9857499	A	17-12-1998	NONE	
WO 9843437	A	01-10-1998	EP 0866616 A	23-09-1998
			WO 9843432 A	01-10-1998
			AU 7038098 A	20-10-1998
			EP 1010331 A	21-06-2000
			NO 994543 A	22-11-1999
			PL 335765 A	22-05-2000
			AU 2701397 A	20-10-1998
			AU 7038198 A	20-10-1998
			WO 9843421 A	01-10-1998
			EP 0866611 A	23-09-1998
			EP 0968610 A	05-01-2000
			EP 1010320 A	21-06-2000
			NO 994534 A	22-11-1999
			NO 994544 A	22-11-1999
			PL 335755 A	22-05-2000
			PL 335780 A	22-05-2000
			AU 2770697 A	20-10-1998
			WO 9843425 A	01-10-1998
			WO 9843426 A	01-10-1998
			WO 9843162 A	01-10-1998
			WO 9843431 A	01-10-1998
			WO 9843248 A	01-10-1998
			WO 9843165 A	01-10-1998
			WO 9843415 A	01-10-1998
			WO 9843172 A	01-10-1998
			WO 9843433 A	01-10-1998
			WO 9843427 A	01-10-1998
			WO 9843167 A	01-10-1998
			WO 9843428 A	01-10-1998
			EP 0872798 A	21-10-1998
			EP 0866613 A	23-09-1998
			EP 0968609 A	05-01-2000
			EP 0968607 A	05-01-2000
			EP 0974229 A	26-01-2000
			EP 0974230 A	26-01-2000
			EP 0968468 A	05-01-2000
			EP 0968465 A	05-01-2000
			EP 0968602 A	05-01-2000

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/JP 00/00135

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9843437 A		EP 0968611 A	05-01-2000
		EP 0968608 A	05-01-2000
		EP 1010068 A	21-06-2000
		EP 0968469 A	05-01-2000
		EP 0972406 A	19-01-2000
		NO 994529 A	19-11-1999
		NO 994530 A	19-11-1999
		NO 994531 A	19-11-1999
		NO 994532 A	19-11-1999
		NO 994533 A	22-11-1999
		NO 994535 A	22-11-1999
		NO 994536 A	22-11-1999
US 5818935 A	06-10-1998	NONE	
EP 0596823 A	11-05-1994	US 5539871 A	23-07-1996
		JP 2677754 B	17-11-1997
		JP 7085243 A	31-03-1995
WO 9800976 A	08-01-1998	US 5903816 A	11-05-1999
		AU 3648797 A	21-01-1998
		EP 0909513 A	21-04-1999
WO 9719560 A	29-05-1997	US 5966120 A	12-10-1999
		AU 7724996 A	11-06-1997
		EP 0862836 A	09-09-1998
		JP 2000500632 T	18-01-2000
WO 9821688 A	22-05-1998	EP 0976089 A	02-02-2000

1287
REPLACED BY
ART 34 AMDT

PATENT COOPERATION TREATY

PCT

REC'D 26 MAR 2001

WIPO

PCT

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

(PCT Article 36 and Rule 70)



14

Applicant's or agent's file reference P/23222.WO/MWM	FOR FURTHER ACTION See Notification of Transmittal of International Preliminary Examination Report (Form PCT/IPEA/416)	
International application No. PCT/IB00/00135	International filing date (day/month/year) 28/01/2000	Priority date (day/month/year) 01/02/1999
International Patent Classification (IPC) or national classification and IPC H04N7/24		
Applicant EMUSE CORPORATION et al.		

1. This international preliminary examination report has been prepared by this International Preliminary Examining Authority and is transmitted to the applicant according to Article 36.
2. This REPORT consists of a total of 6 sheets, including this cover sheet.
☒ This report is also accompanied by ANNEXES, i.e. sheets of the description, claims and/or drawings which have been amended and are the basis for this report and/or sheets containing rectifications made before this Authority (see Rule 70.16 and Section 607 of the Administrative Instructions under the PCT).
These annexes consist of a total of 7 sheets.

3. This report contains indications relating to the following items:

- I ☒ Basis of the report
- II ☐ Priority
- III ☐ Non-establishment of opinion with regard to novelty, inventive step and industrial applicability
- IV ☐ Lack of unity of invention
- V ☒ Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement
- VI ☐ Certain documents cited
- VII ☒ Certain defects in the international application
- VIII ☐ Certain observations on the international application

Date of submission of the demand 30/08/2000	Date of completion of this report 22.03.2001
Name and mailing address of the international preliminary examining authority:  European Patent Office D-80298 Munich Tel. +49 89 2399 - 0 Tx: 523656 epmu d Fax: +49 89 2399 - 4465	Authorized officer Ibruegger, J Telephone No. +49 89 2399 8978 

**INTERNATIONAL PRELIMINARY
EXAMINATION REPORT**

International application No. PCT/IB00/00135

I. Basis of the report

1. This report has been drawn on the basis of *(substitute sheets which have been furnished to the receiving Office in response to an invitation under Article 14 are referred to in this report as "originally filed" and are not annexed to the report since they do not contain amendments (Rules 70.16 and 70.17).):*

Description, pages:

1-21 as originally filed

Claims, No.:

1-28 with telefax of 24/11/2000

Drawings, sheets:

1/11-11/11 as originally filed

2. With regard to the **language**, all the elements marked above were available or furnished to this Authority in the language in which the international application was filed, unless otherwise indicated under this item.

These elements were available or furnished to this Authority in the following language: , which is:

- ☐ the language of a translation furnished for the purposes of the international search (under Rule 23.1(b)).
☐ the language of publication of the international application (under Rule 48.3(b)).
☐ the language of a translation furnished for the purposes of international preliminary examination (under Rule 55.2 and/or 55.3).

3. With regard to any **nucleotide and/or amino acid sequence** disclosed in the international application, the international preliminary examination was carried out on the basis of the sequence listing:

- ☐ contained in the international application in written form.
☐ filed together with the international application in computer readable form.
☐ furnished subsequently to this Authority in written form.
☐ furnished subsequently to this Authority in computer readable form.
☐ The statement that the subsequently furnished written sequence listing does not go beyond the disclosure in the international application as filed has been furnished.
☐ The statement that the information recorded in computer readable form is identical to the written sequence listing has been furnished.

4. The amendments have resulted in the cancellation of:

- ☐ the description, pages:
☒ the claims, Nos.: 29-37

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

International application No. PCT/IB00/00135

☐ the drawings, sheets:

5. ☐ This report has been established as if (some of) the amendments had not been made, since they have been considered to go beyond the disclosure as filed (Rule 70.2(c)):

(Any replacement sheet containing such amendments must be referred to under item 1 and annexed to this report.)

6. Additional observations, if necessary:

V. Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement

1. Statement

Novelty (N)	Yes: Claims 1-28
	No: Claims
Inventive step (IS)	Yes: Claims 1-28
	No: Claims
Industrial applicability (IA)	Yes: Claims 1-28
	No: Claims

2. Citations and explanations
see separate sheet

VII. Certain defects in the international application

The following defects in the form or contents of the international application have been noted:
see separate sheet

**INTERNATIONAL PRELIMINARY
EXAMINATION REPORT - SEPARATE SHEET**

International application No. PCT/IB00/00135

A. ITEM V.

1) The application relates to interactive video programmes.

I. CITED DOCUMENTS

2) The following documents (D) cited in the International Search Report is referred to in this communication; the numbering will be adhered to in the rest of the procedure:

D1 = US-A-5 708 845

II. ARTICLES 33(2) AND 33(3) PCT

3) The application meets the requirements of Articles 33(2) and 33(3) PCT.

a) Claim 1

4) D1 represents the closest prior art from which the subject-matter of the claim is essentially distinguished in that:

- The video programme is parsed to identify separate shots in the video programme to produce an edit list.
- Separate shots in the video programme are identified containing related content to form a sequence of shots containing related content from which a key-frame is used for selecting an object with which data is to be associated.
- Association of interactive data with the object in the sequence of shots.
- Embedding the interactive content data with data representative of said object in a data sequence.

5) Therefore, the subject-matter of claim 1 is new in the sense of Article 33(2) PCT.

6) The features in which the subject-matter of claim 1 is distinguished from D1 provide the advantage of an improved interaction between user and video programme information. In the system of D1 an object in a key-frame is delineated by a user and tracked until a last frame is reached using the outline data. Interactive data are associated with the object by using a hyperlink tool. The interactive data are kept separate from the data representing the object. The claimed subject-matter is not rendered obvious. In particular, there is no suggestion, neither in D1 nor the other documents of the International Search Report, to identify separate shots to produce an edit list and to form a sequence of shots containing related content from which a key-frame is selected.

7) Therefore, the subject-matter of claim 1 involves an inventive step in the sense of Article 33(3) PCT.

b) Claim 14

8) The claim relates to a method underlying the operation of the apparatus specified in claim 1 and essentially recites the features of claim 1 in the language of a method claim.

9) Claim 14 meets the requirements of Article 33(3) PCT for the same reasons as given for claim 1.

c) Claim 27

10) The claim relates to a computer program comprising code for performing all the steps of the method of claim 14.

11) Claim 27 meets the requirements of Article 33(3) PCT in conjunction with claim 14.

d) Dependent claims

12) The dependent claims meet the requirements of Article 33(3) PCT in conjunction with the independent claims to which they refer.

III. ARTICLE 33(4) PCT

13) The claimed subject-matter of the claims is industrially applicable.

B. ITEM VII.

14) Contrary to the requirements of Rule 5.1(a)(ii) PCT, the relevant background art disclosed in the documents D1 is not mentioned in the description, nor is this document identified therein.

15) The description is not in conformity with the claims as required by Rule 5.1(a)(iii) PCT.

16) The features of the claims are not provided with reference signs placed in parentheses (Rule 6.2(b) PCT).

Claims:

1. An interactive system including means for providing a video programme signal, means for generating interactive content data associated with at least one object, said data being associated with frames of said video programme signal in which the object appears, means for multiplexing said data with said video programme signal, means for viewing the video programme signal, means for retrieving said data and means for using said data to obtain details of said object.
2. An interactive system claimed in claim 1, wherein each frame of said video programme includes said interactive content data.
3. An interactive system as claimed in claims 1 or 2, wherein said means for using said data further include means for producing a list of details of said object and means for selecting from said list.
4. An interactive system as claimed in any of claims 1 to 3, wherein said means for using said data include means for accessing an interactive Web site to obtain said details of said object.
5. An interactive system as claimed in claims 3 or 4, wherein said means for accessing an interactive Web site is adapted to secure details of said object which may include a purchasing transaction for said object or browsing an advertising catalogue.
6. An interactive system as claimed in any of the preceding claims, wherein the means for generating includes means for tracking said object in each frame of said video programme signal in which said object appears and means for identifying the location of said object in each said frame.
7. An interactive system as claimed in claim 6, wherein said tracking means includes means for determining scene breaks and

means for searching for said object in a next frame in which said object appears.

8. An interactive system as claimed in any of the preceding claims, wherein said multiplexing means includes means for
5 synchronising said data with audio and video data of said programme signal to generate a transport stream.

9. An interactive system as claimed in claim 8, wherein said
10 system includes means for broadcasting said transport stream via, at least one of a satellite, terrestrial and cable network.

10. An interactive system as claimed in any of the preceding claims, wherein said means for retrieving includes one of a
15 mouse, a keyboard, and remote control device.

11. An apparatus for associating data representative of an object with a digital video programme including means for providing a digital video programme having plural individual
20 frames at least some of which incorporate said object, means for selecting a frame of the video programme in which said object appears to provide a key-frame, means for selecting said object within the key-frame with which data is to be associated, means for extracting attributes of the object from the key-frame, means
25 for associating interactive data with the object in the key-frame, means for utilising the attributes of the object for tracking the object through subsequent frames of the video programme, whereby said interactive data is associated with the object in subsequent frames of the video programme in which said
30 object has been tracked and said interactive content data is embedded with data representative of said object in a data sequence.

12. An apparatus as claimed in claim 11, wherein means are
35 provided for converting said data sequence to a standard data sequence.

13. An apparatus as claimed in claims 11 or 12, including means for converting a video programme in an analogue format to digitised form.

5 14. An apparatus as claimed in any of claims 11 to 13, wherein the means for selecting a frame of the video programme includes means for producing an edit list to divide the digitised video programme into a plurality of sequences of related shots, and means for selecting at least one key-frame from within each
10 sequence.

15 15. An apparatus as claimed in claims 14, wherein the means for producing an edit list further includes means for parsing the video programme by identifying separate shots in the video programme to produce the edit list, means for identifying shots containing related content to form a sequence of shots containing related content, and means for producing a hierarchy of groups of shots.

20 16. An apparatus as claimed in claim 15, wherein said means for parsing include means for inputting criteria to be used to recognise a change of shot.

25 17. An apparatus as claimed in any of claims 11 to 16, wherein the means for extracting attributes of the object includes means for isolating the object within a boundary formed on the frame, means for performing edge detection within the boundary to identify and locate edges of said object, and storing means for storing a geometric model of said object.

30 18. An apparatus as claimed in any of claims 11 to 17, wherein said means for extracting attributes of said object also includes means for recording at least one of the attributes of shape, size, position, colour, texture, intensity gradient of said
35 object, and time series statistics based on said attributes.

19. An apparatus as claimed in any of claims 11 to 18, wherein said means for extracting attributes of said object includes means for comparing said attributes of said object with attributes of objects previously stored to determine whether the object is distinguishable therefrom, and when said object is determined not to be distinguishable, providing means for re-defining the object.

20. An apparatus as claimed in any of claims 11 to 19, wherein said means for extracting said attributes includes means for comparing the location in the frame of said object with the location of objects already stored for that frame to determine whether that object is distinguishable therefrom, and where the location of said object is not distinguishable from the location of another object providing means for assigning rank to the objects to determine which object will be associated with that location.

21. An apparatus, as claimed in any of claims 11 to 20, wherein the means for tracking the object includes means for updating the stored attributes of the object as the object moves location within different frames.

22. An apparatus as claimed in any of claims 11 to 21, wherein said means for tracking includes plural algorithm means for use depending on the visual complexity of a sequence to automatically track objects in different types of visual environment.

23. An apparatus as claimed in any of claims 11 to 22, wherein said tracking means includes means for converting all the frames to be tracked to a low-level representation, means for determining the position of each object in the frames by minimising a distance measure to locate each object in each frame, means for processing the positions of said object to smooth over occlusions and the entrances and exits of objects into and out of said frames, and means for reviewing the object

within a tracked sequence and for correcting the location attributes of any misplaced objects.

24. An apparatus, as claimed in any of claims 11 to 23, wherein the means for associating includes means for providing a database of different types of data including one or more of URLs, HTML pages, video clips, audio clips, text files and multimedia catalogues, and means for selecting said interactive content data from the database to associate with said object.

25. An apparatus, as claimed in any of claims 11 to 24, wherein the means for associating produces said data sequence using means for determining whether the embedded interactive content data is frame synchronous data associated with object positions, shapes, ranks and pointers in a frame, or group-synchronous data associated with all the objects in a group, or is data to be streamed just in time, wherein means are provided for associating frame synchronous data with the corresponding frame, means are provided for associating group synchronous data with the frame at which a group changes, and means are provided for streaming just in time data to a user before it is required to be associated with the corresponding objects.

26. An apparatus as claimed in any of claims 11 to 25, wherein means are provided to associate different interactive content data with respectively different objects.

27. An apparatus for embedding a data sequence within a generic digital transport stream, including means for receiving a data sequence of interactive content data associated with an object in a digitised video signal, means for synchronising the data sequence with the video and audio of the digitised video signal to generate a further transport stream, and means for associating a packet identifier with the further transport stream.

28. An apparatus as claimed in claim 27, wherein means are provided for broadcasting the further transport stream to viewers.

5 29. An apparatus as claimed in claims 27 or 28, wherein the means for receiving a data sequence includes means for receiving elementary streams comprising a digital video signal stream, a digital audio stream, a digital data sequence stream and a digital control data stream, means for packetising each of the
10 data streams into fixed size blocks and adding a protocol header to produce packetised elementary streams, and means for synchronising the packetised elementary streams with time stamps to establish a relationship between the data streams.

15 30. An apparatus as claimed in any of claims 27 to 29, wherein the means for synchronising the data sequence includes means for multiplexing packetised elementary streams into transport packets headed by a synchronisation byte, and means for assigning a different packet identifier to each packetised elementary stream.
20

31. An apparatus as claimed in claim 30, wherein means for synchronising the packetised elementary streams with time stamps includes means for stamping with a reference time stamp to indicate current time, and means for stamping with a decoding
25 time stamp to indicate when the data sequence stream has to be synchronised with the video and audio streams.

32. An apparatus as claimed in claim 28, wherein the means for broadcasting the further transport streams to users includes
30 means for providing a programme association table listing all the channels to be available in the broadcast, means for providing a programme map table identifying all the elementary streams in the broadcast channel, and means for transmitting the programme association table and the programme map table as separate packets
35 within the further transport stream.

33. An apparatus for retrieving data embedded in a generic digital transport stream in which the embedded data includes a data sequence of data associated with objects represented by the generic digital transport stream, said apparatus including means for recognising a packet identifier within the video signal, means for extracting the data sequence from the generic digital transport stream, means for identifying objects within the video sequence from which to retrieve associated data, means for synchronising said data sequence to said identified objects and means for interactively using said associated data.
34. An apparatus as claimed in claim 33, wherein said means for identifying objects includes means for selecting an object within a frame, means for displaying data associated with said object, means for selecting data from a list of displayed data, and means for extracting the embedded data associated with the data relating to said object.
35. An apparatus as claimed in claims 33 or 34, wherein means are provided for selecting a frame to display the objects having embedded associated data, means for selecting one of the displayed objects to display a list of the data associated with said object, and means for selecting from said list.
36. An apparatus as claimed in claim 35, wherein the means for selecting a frame includes means for storing the frame for subsequent display and subsequent recall of the frame.
37. An apparatus as claimed in any of claims 33 to 36, wherein the extracted embedded data is applied to means for accessing an Internet web site to facilitate interactive communication such as e-commerce.

09/889966

IAIN C. BAILLIE B.Sc., I.D., E.P.A., C.P.A., F.I.T.M.A.
MARTYN W. MOLYNEAUX C.Eng., F.I.E.E., E.P.A., C.P.A., M.I.T.M.A.
PETER D. GALLOWAY I.D., E.P.A., C.P.A.
JOHN RICHARDS M.A. Contab., LL.B., E.P.A., C.P.A.

GRAHAM FARRINGTON LL.B., M.I.T.M.A.
CLIFFORD J. WANT C.Eng., M.Sc., M.I.P.T., M.I.N.I.S., C.P.A.
HUGH R. WOTHERSPOON B.Sc., B.A., E.P.A.

52-54 HIGH HOLBORN — LONDON WC1V 6RI

TELEPHONE: 020-7242 5566 FACSIMILE: 020-7405 1908

TELEX: 262433 MUMSEI G B E F T V G 473

e-mail: ip@langnerparry.com

ASSOCIATED OFFICE

DACHAUERSTRASSE 37 — 80335 MUNICH — GERMANY

TELEPHONE: (089) 269077 FACSIMILE: (089) 269040

FACSIMILE

24th November 2000

International Preliminary Examination Authority
The European Patent Office
D-80298 Munich
GERMANY

For the attention of: Mr Ibruegger

Dear Sirs

RE: International Patent Application No. PCT/IB00/00135
EMUSE CORPORATION
Our Ref: P/23222.WO/MWM

Further to our telephone conversation with Mr Ibruegger on 22nd November 2000, we enclose, in duplicate, revised claims 1 to 28 to replace all the claims on file and on which we request that Examination proceed. These amended claims are based on claims 11 and 14 to 26 as filed, in that claims 1 to 13 are based on claims 11 and 14 to 26 as filed, claims 14 to 26 are method claims corresponding to the apparatus claims 1 to 13 filed herewith, and claims 27 and 28 are to a computer program and a computer program product corresponding to the method claims of claims 14 to 26. In particular, claims 1 and 14 filed herewith are based on original claim 11 and features from claims 14 and 15 as filed. Claims 2 and 15 are based on a further feature from claim 15 and claims 3 to 13 correspond to claims 16 to 26 as filed. It is, therefore, submitted that no new matter is added by the new claims

The cancelled claims are deleted without prejudice to the reinstatement of these claims into this application or the incorporation, in the national or regional phase, of said claims in one or more divisional applications, if such action should prove desirable in the future. Moreover, these claims are submitted without prejudice to the subsequent submission of broader claims, within the scope of the application originally filed, should that prove desirable in the future.

As now claimed, claims 1 and 14 are directed to means for, and a method of, parsing a video programme to form sequences of shots containing related content, and then selecting a key-frame within each of the sequence of shots before tracking objects through the frames. This is on the basis a given object is most likely to be found in frames having related content. This has the advantage of avoiding searching all the frames for all the objects selected but confines the search to a sequence or sequences of shots in which the objects are most likely to appear. This has an advantage over the prior art in which it is necessary to search all frames for all objects.

US 7708845 ('845), IBM Technical Disclosure Bulletin, Vol. 37, No. 048, April 1994 and XB-002139169 were cited in the Search Reports in respect of claims 11, 14 and 15 from which new claims 1 and 14 are derived.

IBM Technical Disclosure Bulletin Vol. 37, No. 04A, April 1994, (IBM) refers to a method to link objects from one video to another video or some other form of media. The citation discloses the identification of objects on a first frame in which they appear in the video, means for outlining the object in the frame and refers to the use of a computer image tracking programme to identify the image on future frames, such image tracking programmes being stated to be known for the colouration of black and white movies. Data is then associated with the tracking information. It is, however, submitted that the citation is not an enabling disclosure because no description is given of how the procedure is to be carried out.

XP-002139169 ('169) discloses a method of forming video abstracts for use in retrieving video material from a large video database. The procedure includes dividing a video programme into video segments, possibly corresponding with the video clips of the current invention, in which ends of the segments are determined automatically by detecting a scene cut. Assigning a number of key frames to each segment or clip (paragraph 2.1), a selected object in the key frame is modelled by size, shape or colour distribution (paragraph 3.1 page 917 column 1 last paragraph) and then the object is sought in each new video frame. The model is updated when the object is found in a succeeding frame (see same paragraph). That is, the video stream is segmented automatically into clips where the frame content changes significantly, page 919 column 1 second paragraph. In order to retrieve material from the video database, it is then possible to view clips showing particular objects, page 919 column 1 second paragraph.

There is no disclosure of grouping clips containing the same objects, prior to tracking objects through the frames, as is disclosed in the present invention, on page 11, paragraph 4 to page 15 paragraph 1. That is, there is no disclosure of dividing the video into clips, selecting key frames, and comparing objects within the key frames from different clips and then grouping together clips having key frames containing the same objects before tracking objects through all the frames of the selected clips. In particular, there is no disclosure of means for parsing the video programme by identifying separate shots in the video programme to produce an edit list, means for identifying shots containing related content and means for selecting at least one key frame from within each sequence of shots, as required by amended

claim 1, or of parsing the video programme by identifying separate shots in the video programme to produce an edit list, identifying, from the edit list, shots containing related content to form a sequence of shots containing related content, and selecting at least one key-frame containing the object from within the sequence of shots as required by claim 14 now filed.

The present invention has the advantage of avoiding searching all the frames for all the objects selected but confines the search to the clips in which the object is known, or expected, to appear. On the contrary, in '169, objects are tracked through successive frames until a frame is reached in which the object does not appear. Tracking then ceases for that object, even although the object may re-appear in later frames.

US 5708845 ('845) discloses use of frame data and object mapping data in which the two types of data are kept separately, column 2 lines 41 to 58. Also disclosed is the outlining of an object in a key frame, column 11 line 12, then using a motion tracking tool to track the object through subsequent frames. Data associated with the object may be stored for the key frame and the last consecutive frame in which the object appears.

'845 further discloses a means of adding interactive data to a video stream in which the frame data and the interactive data are stored in separate files. That is, the object data is not embedded in the data stream, as required by claims 1 and 14. Having selected an object within a frame, the object may then be tracked through subsequent frames using an object motion tracking tool, column 3 lines 18 to 23. As an alternative to a user tracking objects through the frames, a list of frames in which the object appears may be supplied by the video owner, column 7 lines 20 to 25. However, no disclosure is given as to how the owner determines in which frames the object appears. In order to track objects through the frames, an object is first outlined on a given frame, column 9 line 66 to column 10 line 2 and this process is repeated for all objects to be mapped in a frame and for all frames, column 10, lines 10 to 12. In order to provide some data compression, rather than indexing every frame, it is suggested that a first and last frame of the sequence of frames in which an object appears be indexed, column 10 lines 20 to 25. Similarly, a motion tracking tool may be used for tracking a moving object from frame to frame, column 10 lines 28 to 30. The object is tracked from frame to frame, even where there may be a rotation or occlusion of the object, as long as it retains some recognised features, column 10 lines 51 to 53. Similarly, MPEG-2 techniques can be used for separating objects from a stationary background, column 10 lines 58 to 65, that is, a moving object can be tracked by marking its position in a key frame, column 11 lines 10 to 12 and then detecting the image across subsequent frames, column 11 lines 15 to 18. For an object having a regular motion, some compression can be obtained by marking a position of the object at every N frames and interpolating, column 11 lines 56 to 65.

'845, therefore, does not disclose the initial dividing up of the video into shots containing the same objects and the grouping of shots containing the same objects before searching frames from objects, as disclosed in the present invention.

In particular, there is no disclosure of means for parsing the video programme by identifying separate shots in the video programme to produce an edit list, means for identifying shots containing related content and means for selecting at least one key frame from within each sequence of shots, as required by amended claim 1, or of parsing the video programme by identifying separate shots in the video programme to produce an edit list, identifying, from the edit list, shots containing related content to form a sequence of shots containing related content, and selecting at least one key-frame containing the object from within the sequence of shots as required by claim 14 now filed.

Moreover, no combination of '169, '845 and IBM would, we submit, lead to the parsing feature of this invention defined in the independent claims now filed.

'845 was also cited in relation to original claim 17 from which new claims 4 and 17 are derived.

Amended claims 4 and 17 define apparatus for, and method of, respectively, detecting an object by locating an edge within a boundary and storing a geometric model of the object for tracking in subsequent frames. US 7508845 ('845) discloses in column 9 lines 44 to 45, the outlining of a shape and storing of shape data, and at lines 50 to 54 the outlining of a object with a cursor and storing co-ordinates of the shape, and at column 9 lines 67 to column 10 line 2 there is a disclosure of drawing an outline around an object, where the outline may be equivalent to the boundary of the current application, and then the co-ordinates of the outline are saved, column 10 line 4. In column 11 lines 30 onwards, there is disclosed a drawing of an outline round an object and marking of the central position of the object, and the use of a motion tracking tool to then track the object from frame to frame. In column 10 lines 40 to 41 there is a disclosure of feature segmentation and clustering to form an abstracted cluster, representative of objects. There is no disclosure in '845 of drawing a boundary around an object and then detecting the edges of the object within the boundary. IBM Technical Disclosure Bulletin 37 (04A) April 1994 (IBM), discloses a use of a mouse etc to outline an object, in the second paragraph 4th sentence. Once again, however, there is no disclosure of detecting edges of an object within such an outline. It is submitted, therefore, that the features of claims 4 and 17 are novel and inventive with respect to these references.

'845 and IBM have been cited against original claim 18 on which new claims 5 and 18 are based.

Amended claims 5 and 18 define extracting and recording attributes of objects such as size, position, colour, texture and intensity gradient, and time series statistics based on these attributes as well as shapes. US 5708845 ('845) discloses in column 9 line 44 to 45 only the storage of a shape of an object and in column 10 lines 40 to 41, discloses feature segmentation and clustering techniques to produce an abstracted cluster representation of objects. Claims 5 and 18 therefore include other attributes, and since these claims are appended to independent claims which we submit have novelty and inventive step, then the dependent claims have novelty and inventive step.

IBM discloses, in second paragraph, seventh sentence, the saving of a relative form of tracking information regarding the identified object. There is, therefore, no disclosure in either of these cited references of the recording of the specified attributes, apart from shape, and claims 5 and 18 therefore add further novel features with respect to the cited references.

'845 has been cited against claim 19 on which new claims 6 and 19 are based.

Amended claims 6 and 19 define the feature of redefining an object when a new object is not distinguishable from previously stored objects. The Search Examiner has indicated that this is not novel with respect to US 5708845 ('845). '845 discloses the mapping of objects in column 9 to column 12, but there is no disclosure of the comparing of new objects with objects already stored and redefining the objects if they are not distinguishable from those already stored. Claims 6 and 19 therefore, add further novel features, in the light of the cited reference.

'845 and EP 0675461A have been cited in combination against original claim 20 on which new claims 7 and 20 are based.

Amended claims 7 and 20, define the feature of comparing the location of an object with the locations of objects already stored for a given frame and to assigning rank, if the objects are not distinguishable. As already discussed, US 5708845 ('845) discloses a tracking of objects from frame to frame. EP 0675461A ('461) discloses a method and apparatus for producing animations. Rather than storing the colour of each pixel of a frame, '461 discloses the storing of outlines and the colour associated with an outline, so that the area within the outline can be coloured with the stored colour. In columns 3 lines 24 to 40, column 18 lines 28 to 33, column 19 lines 15 to 21 and column 21 lines 15 to 29, there is disclosed a procedure for dealing with the colouring of overlapping objects in which priority is assigned to the boundary lines of the objects, column 19 lines 15 to 21 to determine which colour should be used. We submit, the person skilled in the art would not look to the field of animation for determining how to deal with the problem of overlapping objects being tracked on a video. In any case, the procedure of moving objects towards the back or front within drawing programs is well known, however, we submit that this does not equate to assigning a rank to an object with particular co-ordinates on the screen when more than one object occupies those co-ordinates. Therefore, a combination of the cited references does not lead to the feature of claims 7 and 20.

Claims 8 and 21 define the feature of updating object attributes from frame to frame, which the Search Examiner has indicated in respect of original claim 21, on which these claims are based, is not novel in the light of US 5708845 ('845). This citation discloses in column 10 lines 50 to 55 that an object may be recognised from frame to frame if it retains some recognised features. However, there is no disclosure of changing the object attributes and there is no disclosure in this citation of updating object attributes as the object is tracked from frame to frame and

therefore it is submitted claims 9 and 21 add novel and inventive features in respect of this citation.

'845 was cited with references to original claim 22 on which new claim 9 and 22 are based.

Claims 9 and 22 include the feature for "calculation of independent tracks of objects". This has basis in figure 7 and the description on page 15 lines 21 to 22 where it is indicated that the objects first have data embedded and then all the objects are tracked through successive frames. This feature is not disclosed in '845.

Claims 10 and 23 define the features of 1) converting frames to a low-level representation, 2) determining the position of objects by minimising a distance measure, 3) processing the positions to smooth over occlusions, exits and entrances etc and 4) correcting the location of misplaced objects. The Search Examiner has indicated that original claim 23, on which these claims are based, is not inventive in the light of XP-002139169 ('169) and US 5708845 ('845). '169 discloses segmenting a video into a static background and moving objects in order to form an abstract of a video for subsequent retrieval of videos of interest from a video database. The citation makes reference, on page 916, section 2.3, third paragraph to the use of parameters that relate tracked objects to individual video frames but there is no disclosure of which these parameters are. On page 917, paragraph bridging columns 1 and 2, it is disclosed that an object is modelled by its position and velocity in a stabilised scene, however, there is no disclosure of determining a position of objects by minimising a distance measurement as required by claims 10 and 23. Therefore, '169 does not appear to disclose any of the features of claims 10 and 23.

US 5708845 ('845) discloses the storing of display location co-ordinates of objects intended to be interactive as they appear in display frames, column 2 lines 45 to 50. Means of defining the location of an object are described in column 5 lines 55 to 67. However, we can locate no disclosure in '845 to the specific features of claims 10 and 23.

Therefore, we submit that the combination of '169 and '845 does not lead to the features of claims 10 and 23.

Claims 11 and 24 define features of providing a database of different types of data including URLs, HTML pages, video clips, audio clips, text files and multimedia catalogues from which interactive content data may be selected to associate with an object. The Search Examiner has indicated that original claim 24, on which these claims are based, is not novel in the light of US 5708845 ('845). This citation discloses in column 2, lines 50 to 55, linkages associated with the objects to respective other functions to be performed upon user selection of the objects. In column 5 lines 10 to 15, it is indicated that in response to a selection of an object, an interactive digital media program responds by launching further layers of display presentations and/or triggering other program functions, such as launching another application, initiating the operation of another system, or

connecting to an external network such as a World Wide Web page or service on the Internet. In column 8 line 55 to column 9 line 3, it is indicated that on selecting an object, a pop-up window, overlay display, or audio track may be presented, or another program may be executed such as initiating an Internet connection for on-line purchasing. In column 13 lines 15 to 20, it is indicated that a text file may be displayed on selecting an object. However, there is no disclosure of the provision of a database as required by claims 11 and 24.

Claims 12 and 25 define a means and method respectively, of determining whether the embedded interactive data is frame synchronous data associated with an object or group-synchronous data associated with all the objects in a group, or is data to be streamed just in time, and corresponding means, or method steps, are provided for associating the data. The Search Examiner has indicated that original claim 25, on which these claims are based, was not novel in the light of US 5708845 ('845). In column 2 lines 50 to 55 the citation discloses the association of interactive data with objects. The citation also discloses that the display location coordinates of objects and the frame addresses of the frames are stored separately, column 5 line 64 to 67 and column 6 lines 1 to 21. Since the frame data and the interactive data are stored separately, it may be suggested that this an implementation of "just in time" association of the frames, but there is no specific disclosure of transmitting the interactive data just in time before it is required to be associated with corresponding objects as required by claims 12 and 25, on which these claims. We are not aware of any disclosure in '845 of the association of interactive data with all the objects in a group. We submit that claims 12 and 25 are, therefore, novel and inventive with respect to the cited reference.

We further submit that since claims 2 to 13 and 15 to 26 are appended directly or indirectly to independent claims which we submit are novel and inventive, then the independent claims, it is submitted, are novel and inventive.

We submit that, as revised, the new claims define patentable matter and the Examiner's favourable consideration is requested.

We will revise the statements of invention in conformity with the revised claims when the Examiner has agreed patentable claims.

We look forward to receiving a favourable report from the Examiner and thank the Examiner for his assistance.

Yours faithfully
LANGNER PARRY

CLIFFORD J WANT

CJW/kj

Claims:

1. An apparatus for associating data representative of an object with a digital video programme including means for providing a digital video programme having plural individual frames at least some of which incorporate said object, means for parsing the video programme by identifying separate shots in the video programme to produce an edit list, means for identifying shots containing related content to form a sequence of shots containing related content, means for selecting at least one key-frame from within each sequence of shots, means for selecting said object within the key-frame with which data is to be associated, means for extracting attributes of the object from the key-frame, means for associating interactive data with the object in the key-frame, tracking means for utilising the attributes of the object for tracking the object through the sequence of shots, whereby said interactive data is associated with the object in the sequence of shots and said interactive content data is embedded with data representative of said object in a data sequence.
2. An apparatus as claimed in claim 1, wherein the means for identifying shots containing related content to form a sequence of shots containing related content includes means for producing a hierarchy of groups of shots.
3. An apparatus as claimed in claims 1 or 2, wherein said means for parsing include means for inputting criteria to be used to recognise a change of shot.
4. An apparatus as claimed in any of claims 1 to 3, wherein the means for extracting attributes of the object includes means for isolating the object within a boundary formed on the frame, means for performing edge detection within the boundary to identify and locate edges of said

object, and storing means for storing a geometric model of said object.

5. An apparatus as claimed in any of claims 1 to 4, wherein said means for extracting attributes of said object also includes means for recording at least one of the attributes of shape, size, position, colour, texture, intensity gradient of said object, and time series statistics based on said attributes.

6. An apparatus as claimed in any of the preceding claims, wherein said means for extracting attributes of said object includes means for comparing said attributes of said object with attributes of objects previously stored to determine whether the object is distinguishable therefrom, and when said object is determined not to be distinguishable, providing means for re-defining the object.

7. An apparatus as claimed in any of the preceding claims, wherein said means for extracting attributes of said object includes means for comparing the location in the frame of said object with the location of objects already stored for that frame to determine whether that object is distinguishable therefrom, and where the location of said object is not distinguishable from the location of another object providing means for assigning rank to the objects to determine which object will be associated with that location.

8. An apparatus, as claimed in any of the preceding claims, wherein the means for utilising the attributes of the object for tracking the object includes means for updating the stored attributes of the object as the attributes of the object change from frame to frame.

9. An apparatus as claimed in any of the preceding claims, wherein said tracking means utilising the attributes of the object for tracking the object includes

plural algorithm means for calculation of independent tracks of objects for use depending on the visual complexity of a sequence to automatically track said objects in different types of visual environment.

10. An apparatus as claimed in any of the preceding claims, wherein said tracking means for utilising the attributes of the object for tracking the object includes means for converting all the frames to be tracked to a low-level representation, means for determining the position of each object in the frames by minimising a distance measure to locate each object in each frame, means for processing the positions of said object to smooth over occlusions and the entrances and exits of objects into and out of said frames, and means for reviewing the object within a tracked sequence and for correcting the location attributes of any misplaced objects.

11. An apparatus, as claimed in any of the preceding claims, wherein the means for associating interactive data with the object in the key-frame includes means for providing a database of different types of data including one or more of URLs, HTML pages, video clips, audio clips, text files and multimedia catalogues, and means for selecting said interactive content data from the database to associate with said object.

12. An apparatus, as claimed in any of the preceding claims, wherein the means for associating interactive data with the object in the key-frame produces said data sequence using means for determining whether the embedded interactive content data is frame synchronous data associated with object positions, shapes, ranks and pointers in a frame, or group-synchronous data associated with all the objects in a group, or is data to be streamed just in time, wherein means are provided for associating frame synchronous data with the corresponding frame, means are provided for associating group

synchronous data with the frame at which a group changes, and means are provided for streaming just in time data to a user before it is required to be associated with the corresponding objects.

13. An apparatus as claimed in any of the preceding claims, wherein means are provided to associate different interactive content data with respectively different objects.

14. A method for associating interactive data representative of an object with a digital video programme including the steps of:

- a) providing a digital video programme having a plurality of individual frames at least some of which incorporate said object with which data is to be associated,
- b) parsing the video programme by identifying separate shots in the video programme to produce an edit list,
- c) identifying, from the edit list, shots containing related content to form a sequence of shots containing related content,
- d) selecting at least one key-frame containing the object from within the sequence of shots,
- e) locating said object within the at least one key-frame,
- f) extracting attributes of the object from the at least one key-frame,
- g) associating interactive data with the object in the at least one key-frame,
- h) tracking the object through the sequence of shots utilising the attributes of the object,
- i) associating said interactive data with the object in frames in the sequence of shots, and
- j) embedding said interactive data with data representative of said object in a data sequence representative of the digital video programme.

15. A method as claimed in claim 14, wherein step b) includes the step of inputting criteria to be used to recognise a change of shot.
16. A method as claimed in claims 14 or 15, wherein step c) includes the step of producing a hierarchy of groups of sequences of shots.
17. A method as claimed in any of claims 14 to 16, wherein step e) includes the steps of: isolating the object within a boundary formed on the frame, performing edge detection within the boundary to identify and locate edges of said object, and step f) includes storing a geometric model of said object.
18. A method as claimed in any of claims 14 to 17, wherein step f) includes the step of recording at least one of the attributes of shape, size, position, colour, texture, intensity gradient of said object, and time series statistics based on said attributes.
19. A method as claimed in any of claims 14 to 18, wherein step f) includes the step of comparing said attributes of said object with attributes of objects previously stored to determine whether the object is distinguishable therefrom, and when said object is determined not to be distinguishable, the step of re-defining the object.
20. A method as claimed in any of claims 14 to 19, wherein step f) includes the step of comparing the location in the frame of said object with the location of objects already stored for that frame to determine whether that object is distinguishable therefrom, and where the location of said object is not distinguishable from the location of another object, the step of assigning rank to the objects to determine which object will be associated with that location.

21. A method, as claimed in any of claims 14 to 20, wherein step h) includes the step of updating the stored attributes of the object as the attributes of the object change from frame to frame.
22. A method as claimed in any of claims 14 to 21, wherein step h) includes the step of using plural algorithm means for calculation of independent tracks of objects for use depending on the visual complexity of a sequence automatically to track said objects in different types of visual environment.
23. A method as claimed in any of claims 14 to 22, wherein step h) includes the steps of converting all the frames to be tracked to a low-level representation, determining the position of each object in the frames by minimising a distance measure to locate each object in each frame, processing the positions of said object to smooth over occlusions and the entrances and exits of objects into and out of said frames, reviewing the object within a tracked sequence and correcting the location attributes of any misplaced objects.
24. A method, as claimed in any of claims 14 to 23, wherein step g) includes the steps of providing a database of different types of data including one or more of URLs, HTML pages, video clips, audio clips, text files and multimedia catalogues, and selecting said interactive content data from the database to associate with said object.
25. A method, as claimed in any of claims 14 to 24, wherein step j) includes determining whether the embedded interactive content data is frame synchronous data associated with object positions, shapes, ranks and pointers in a frame, or group-synchronous data associated with all the objects in a group, or is data to be streamed just in time, and associating frame synchronous data with the corresponding frame, associating group

synchronous data with the frame at which a group changes, and streaming just in time data to a user before it is required to be associated with the corresponding objects, respectively.

26. A method as claimed in any of claims 14 to 25, wherein in steps d) to j) different interactive content data are associated with respectively different objects.

27. A computer program comprising code means for performing all the steps of the method of any of claims 14 to 26 when the program is run on one or more computers.

28. A computer program as claimed in claim 27, wherein the computer program is embodied on a computer-readable medium.



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁷ :

H04N 7/24

A2

(11) International Publication Number:

WO 00/45599

(43) International Publication Date:

3 August 2000 (03.08.00)

(21) International Application Number: PCT/IB00/00135

(22) International Filing Date: 28 January 2000 (28.01.00)

(30) Priority Data:

9902235.2

1 February 1999 (01.02.99)

GB

(71) Applicant (for all designated States except US): EMUSE CORPORATION [IE/IE]; 80 Park Avenue, Sandymount, Dublin 4 (IE).

(72) Inventor; and

(75) Inventor/Applicant (for US only): RAINSFORD, Patrick, J. [IE/IE]; 80 Park Avenue, Sandymount, Dublin 4 (IE).

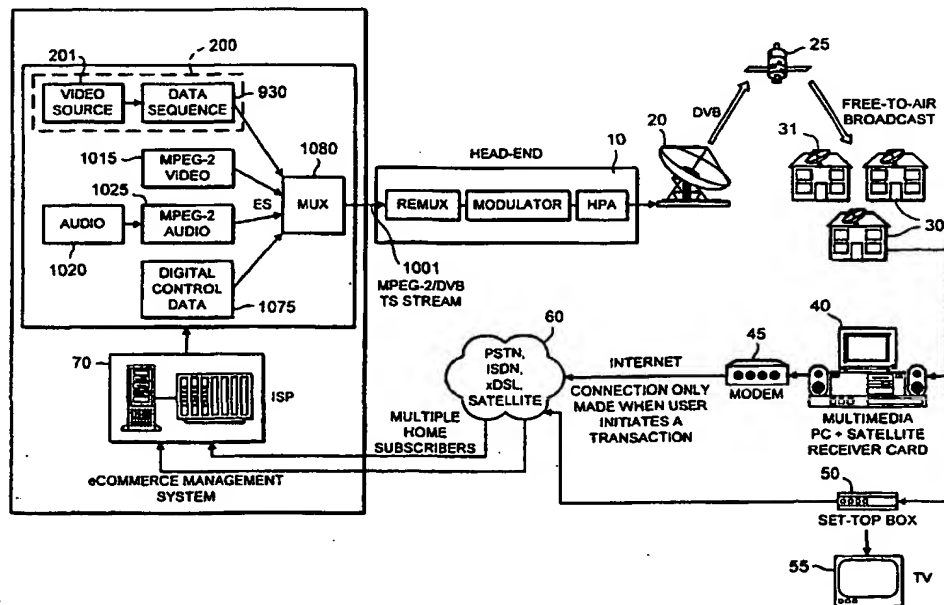
(74) Agents: BAILLIE, IAIN, C. et al.; Langner Parry, 52-54 High Holborn, London WC1V 6RR (GB).

(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published

Without international search report and to be republished upon receipt of that report.

(54) Title: INTERACTIVE SYSTEM



(57) Abstract

An interactive system provides a video programme signal and generates interactive content data to be associated with at least one object within a frame of the video programme. The interactive content data is embedded with the object and the object is tracked through a sequence of frames and the interactive content data is embedded into each one of the frames. The programme frames with the embedded data are multiplexed with video and audio signals and may be broadcast. A receiver identifies an object of interest and the embedded data associated with the object is retrieved. The embedded data may be used for e-commerce.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Larvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon	KR	Republic of Korea	PL	Poland		
CN	China	KZ	Kazakstan	PT	Portugal		
CU	Cuba	LC	Saint Lucia	RO	Romania		
CZ	Czech Republic	LI	Liechtenstein	RU	Russian Federation		
DE	Germany	LK	Sri Lanka	SD	Sudan		
DK	Denmark	LR	Liberia	SE	Sweden		
EE	Estonia			SG	Singapore		

INTERACTIVE SYSTEM

This invention relates to an interactive system and particularly to a system for multiplexing data in a digital video signal.

It is known to provide a video programme in the form of a digital signal which may be broadcast, or which may be provided on a digital video disk (DVD) or a video tape and the present invention is not restricted to the form in which the video signal for a programme is provided.

With the increasing number of television broadcasting channels, there is a dilution of advertising revenue since, for commercial reasons, an advertiser restricts their marketing effort to a limited number of broadcast channels. In addition, there is an increase in availability of devices available to a viewer for preventing the reception of unwanted advertisements, e.g. a V-chip, but at the present time there is currently no way of selectively blocking advertisements, with the result that those advertisements that may be of interest to a viewer are also blocked.

With the growing use of the Internet, users are becoming accustomed to having access to large and diverse sources of data and information using a personal computer (PC) or, for example, a digital set-top box used in conjunction with a television and remote control or mouse.

The present invention seeks to provide a system which enables a viewer to interact with a video signal which may be broadcast so as to facilitate information transfer and/or transactions that may be performed over the Internet.

According to one aspect of this invention there is provided an interactive system including means for providing a video programme signal, means for generating interactive content data associated with at least one object, said data being associated with frames of said video programme signal in which the object appears, means for multiplexing said data with said video programme signal, means for viewing the video programme signal,

means for retrieving said data and means for using said data to obtain details of said object.

Preferably, said means for using include means for accessing an interactive Web site to obtain said details of said object. Conveniently, said means for using further include means for producing a list of details of said object and means for selecting from said list.

Advantageously, said means for accessing an interactive Web site is adapted to secure details of said object which may include a purchasing transaction for said object or browsing an advertising catalogue.

Preferably, the means for generating includes means for tracking said object in each frame of said video programme signal in which said object appears and means for identifying the location of said object in each said frame.

Preferably, each frame of said video programme includes said interactive content data.

Advantageously, said tracking means includes means for determining scene breaks and means for searching for said object in a next frame in which said object appears.

Conveniently, said multiplexing means includes means for synchronising said data with audio and video data of said programme signal to generate a transport stream, for example, a MPEG-2/DVB transport stream.

Advantageously, said system includes means for broadcasting said transport stream via, for example, at least one of a satellite, terrestrial and cable network.

Conveniently, said means for retrieving includes one of a mouse, a keyboard, and remote control device.

According to a second aspect of this invention there is provided apparatus for associating data representative of an object with a digital video programme including means for providing a digital video programme having plural individual frames at least some of which incorporate said object, means for selecting a frame of the video programme in which said object appears to provide a key-frame, means for selecting said object within the key-frame with which data is to be associated, means

for extracting attributes of the object from the key-frame, means for associating interactive data with the object in the key-frame, means for utilising the attributes of the object for tracking the object through subsequent frames of the video programme, whereby said interactive data is associated with the object in subsequent frames of the video programme in which said object has been tracked and said interactive content data is embedded with data representative of said object in a data sequence.

Advantageously, means are provided for converting said data sequence to a standard data sequence, for example, an MPEG-2/DVB compliant data sequence.

Where the video programme is in an analogue format means are preferably provided for converting the programme to digitised form.

Preferably, the means for selecting a frame of the video programme includes means for producing an edit list to divide the digitised video programme into a plurality of sequences of related shots, and means for selecting at least one key-frame from within each sequence.

Advantageously, the means for producing an edit list further includes means for parsing the video programme by identifying separate shots in the video programme to produce the edit list, means for identifying shots containing related content to form a sequence of shots containing related content, and means for producing a hierarchy of groups of shots.

Advantageously, said means for parsing include means for inputting criteria to be used to recognise a change of shot.

Preferably, the means for extracting attributes of the object includes means for isolating the object within a boundary formed on the frame, means for performing edge detection within the boundary to identify and locate edges of said object, and storing means for storing a geometric model of said object.

Conveniently, said means for extracting attributes of said object also includes means for recording at least one of the attributes of shape, size, position, colour, texture, intensity

gradient of said object, and time series statistics based on said attributes.

Advantageously, said means for extracting attributes of said object includes means for comparing said attributes of said object with attributes of objects previously stored to determine whether the object is distinguishable therefrom, and when said object is determined not to be distinguishable, providing means for re-defining the object, for example by re-defining said boundary.

Preferably, said means for extracting said attributes includes means for comparing the location in the frame of said object with the location of objects already stored for that frame to determine whether that object is distinguishable therefrom, and where the location of said object is not distinguishable from the location of another object providing means for assigning rank to the objects to determine which object will be associated with that location.

Preferably, the means for tracking the object includes means for updating the stored attributes of the object as the object moves location within different frames.

Advantageously, said means for tracking includes plural algorithm means for use depending on the visual complexity of a sequence to automatically track objects in different types of visual environment.

Advantageously, said tracking means includes means for converting all the frames to be tracked to a low-level representation, means for determining the position of each object in the frames by minimising a distance measure to locate each object in each frame, means for processing the positions of said object to smooth over occlusions and the entrances and exits of objects into and out of said frames, and means for reviewing the object within a tracked sequence and for correcting the location attributes of any misplaced objects.

Preferably, the means for associating includes means for providing a database of different types of data including one or more of URLs, HTML pages, video clips, audio clips, text files and multimedia catalogues, and means for selecting said

interactive content data from the database to associate with said object.

Preferably, the means for associating produces said data sequence using means for determining whether the embedded interactive content data is frame synchronous data associated with object positions, shapes, ranks and pointers in a frame, or group-synchronous data associated with all the objects in a group, or is data to be streamed just in time, wherein means are provided for associating frame synchronous data with the corresponding frame, means are provided for associating group synchronous data with the frame at which a group changes, and means are provided for streaming just in time data to a user before it is required to be associated with the corresponding objects.

It will be understood that although the above has been defined in relation to associating interactive content data with one object, different interactive content data may be associated with respectively different objects.

According to a third aspect of this invention there is provided apparatus for embedding a data sequence within a generic digital transport stream (such as DVB/MPEG-2 or ATSC/MPEG-2) including means for receiving a data sequence of interactive content data associated with an object in a digitised video signal, means for synchronising the data sequence with the video and audio of the digitised video signal to generate a further transport stream, and means for associating a packet identifier with the further transport stream.

In a preferred embodiment, means are provided for broadcasting the further transport stream to viewers.

Preferably, the means for receiving a data sequence includes means for receiving elementary streams comprising a digital video signal stream, a digital audio stream, a digital data sequence stream and a digital control data stream, means for packetising each of the data streams into fixed size blocks and adding a protocol header to produce packetised elementary streams, and means for synchronising the packetised elementary streams with time stamps to establish a relationship between the data streams.

Preferably, the means for synchronising the data sequence includes means for multiplexing packetised elementary streams into transport packets headed by a synchronisation byte, and means for assigning a different packet identifier to each packetised elementary stream.

Advantageously, means for synchronising the packetised elementary streams with time stamps includes means for stamping with a reference time stamp to indicate current time, and means for stamping with a decoding time stamp to indicate when the data sequence stream has to be synchronised with the video and audio streams.

Conveniently, the means for broadcasting the further transport stream to users includes means for providing a programme association table listing all the channels to be available in the broadcast, means for providing a programme map table identifying all the elementary streams in the broadcast channel, and means for transmitting the programme association table and the programme map table as separate packets within the further transport stream.

According to a fourth aspect of this invention there is provided apparatus for retrieving data embedded in a generic digital transport stream in which the embedded data includes a data sequence of data associated with objects represented by the generic digital transport stream, said apparatus including means for recognising a packet identifier within the video signal, means for extracting the data sequence from the generic digital transport stream, means for identifying objects within the video sequence from which to retrieve associated data, means for synchronising said data sequence to said identified objects and means for interactively using said associated data.

Preferably, said means for identifying objects includes means for selecting an object within a frame, means for displaying data associated with said object, means for selecting data from a list of displayed data, and means for extracting the embedded data associated with the data relating to said object.

Conveniently, means are provided for selecting a frame to display the objects having embedded associated data, means for

selecting one of the displayed objects to display a list of the data associated with said object, and means for selecting from said list.

Conveniently, the means for selecting includes means for storing the frame for subsequent display and subsequent recall of the frame.

In a preferred embodiment, the extracted embedded data is applied to means for accessing an Internet web site to facilitate interactive communication such as e-commerce.

By using the present invention, advertisements produced by advertisers are unobtrusive, i.e. the viewer can watch the programme without interacting, if so desired. Alternatively, the viewer can view the programme and freeze a frame of the programme, click on an object using a mouse, keyboard or TV remote control and, over the Internet, facilitate an e-commerce transaction. In performing such a function the viewer may split the VDU screen so that one portion continues to display the running programme and another portion displays the frozen frame and the Internet information transfer.

The invention can be used in numerous aspects of digital video entertainment, especially broadcasting, i.e.

1. Interactive product placement in regular television programmes or movies.
2. Fashion TV.
3. Music TV.
4. Educational programmes.

The e-commerce may facilitate, for example, merchandising to ticket sales.

The invention has the advantage that a viewer is able to select further information on those items of interest within a video signal programme without being overwhelmed with information of no relevance. This is particularly useful where the information is in the form of advertisements and is achieved by making objects viewed in the video programme have associated multiplexed (embedded) data to provide links to further information relevant to those objects, either to information

within the video signal or stored in a database or by accessing an Internet web site.

As far as the advertiser is concerned, the invention has the advantage that advertisements can be precisely targeted to a relevant audience and the advertisements cannot be stopped from reaching the user by a device for blocking out advertisements, e.g. a V-chip. Because multiple advertisers may associate their advertisements with each frame of a video programme sequence, the invention has the potential of reducing the costs of advertising to individual advertisers while maintaining or increasing advertising revenues for programme makers and suppliers. In this way, data-carrying potential of each frame of a video programme signal may be maximised and maximum use of the data-carrying capacity of broadcast channels may be achieved. The present invention is believed to lead the way to generating a new democracy for advertisers that may not be able to afford, for example, a two minute segment on broadcast TV at peak times. This is because the present invention allows multiple advertisers per object, and/or multiple objects per frame, leading to a high level of flexibility in advertising revenue models.

In the field of, for example, music videos, the content may be used to promote the music of the band for the record label and by interacting with the musicians, a user may purchase and download the music directly.

Additionally, plural advertisers may be buying the same slot - in other words, the advertiser's content is totally fused within the programme content and it is not until the advertising content is downloaded by the user that it is read. Thus, every frame of a digital TV programme may be used as advertising revenue. An e-commerce database may store all relevant data concerning the advertisers, from URL addresses of Web sites to catalogues, brochures and video promotions, to e-commerce transaction facilities.

When a viewer selects an object by, for example, using a mouse to click on the object, that object may represent a number of advertisers, e.g. a musician may advertise clothing, a watch, cosmetics, and a musical instrument, so that the viewer selects

from a list of promoted items associated with the object. There is, thus, presented a push technology approach which maximises the transmission speed of a satellite broadcast. The user needs only a return path via the Internet if he actually wishes to carry out a transaction.

The invention will now be described, by way of example, with reference to the accompanying drawings, in which:

Figure 1 shows a block schematic diagram of an interactive system of this invention,

Figure 2 shows a block schematic diagram of video programme processing for generating interactive content data associated with an object in relevant frames of a programme,

Figure 3 shows a schematic diagram indicating programme sequences derived by groups of related camera shots,

Figure 4 shows a block schematic diagram of a parser shown in Figure 2, whereby groups of shots are produced,

Figure 5 shows a key frame of a video programme,

Figure 6 shows an object selected in the key frame of Figure 5,

Figure 7 shows a flow diagram for frame by frame identification of objects in a video programme,

Figure 8 shows a flow diagram of the object tracker shown in Figure 2 for tracking the object frame by frame,

Figure 9 shows a flow diagram of the streamer shown in Figure 2,

Figure 10 shows a block schematic diagram for combining the interactive content data with the video programme signal,

Figure 11 shows the structure of a data packet used in this invention, and

Figure 12 shows in block schematic form the manner of extracting the interactive content data from the video programme signal.

In the Figures like reference numerals denote like parts.

The interactive system shown in Figure 1 has apparatus 200 for producing a data sequence that is representative of interactive content data associated with at least one object which is multiplexed 1080 with video and audio data

representative of the digital video programme. In the described embodiment, a data transport stream 1001 is applied to head end apparatus 10 of a satellite broadcast device 20 that transmits to a satellite 25 that, in turn, re-transmits the broadcast signal to plural users/viewers 30 each having a respective broadcast receiving dish 31. The received signal may be applied to a PC 40 having a TV card for interaction by a viewer. The received broadcast signal may also, or alternatively, be applied to a set top box 50 of a digital television 55 or a television with integrated set top box electronics. The set top box may be provided with a keyboard (not shown) or a mouse 56 for a viewer to manipulate an icon on the TV to select objects and interact with menus and operations that may be provided. The PC 40 may similarly be provided with a keyboard, but, as is customary, also a mouse so that the manner of use is the same as the set top box, so a viewer/user is able to select an object and perform interactive communication. Input and output to and from the PC is via a modem 45 to a public telephone network 60 which may be, for example, PSTN, ISTN, xDSL, or satellite, and the set top box 50 is similarly connected to the network 60. The network 60 interconnects the multiple viewers with an e-commerce management system 70 that may be a dedicated management system or a system inter-linked with an Internet service provider. In a system where a video programme is broadcast, the system 70 is connected to the broadcast providing system so that the system 70 can tie-in with the broadcast programme for maintaining a reference between the objects transmitted to a viewer.

In the system of this invention an object which may be, for example, a person, physical objects such as clothing, a watch, cosmetics, musical instruments or, for example, a trademark has data associated with that object multiplexed (embedded) into the video programme signal of the programme that carries the object. To achieve this it is necessary to identify and track objects frame by frame throughout the video programme. It is to be understood that although in the described embodiment the video programme is broadcast, the video programme could be on a digital video disk (DVD), tape or any known means for storing a video

programme. The viewer upon selecting an object is then able to interact with details concerning the object. For example, where the object is a musician in a pop musical video, information may be derived as to where the music record, clothing worn and advertised by the musician may be secured over the Internet.

The first stage is to produce the interactive data that will be dynamically associated with the, or each, object in every frame of a programme in which the object appears. A five-minute video sequence, for example, will typically consist of 7,500 frames, whereas a ninety-minute movie may be 135,000 frames.

If the input video programme is not in a digital format, the programme must first be digitised by means known per se.

Referring to Figure 2, the apparatus 200 for generating the interactive content data associated with an object in relevant frames of a programme is shown. The digitised programme from a digital video source 201 is divided into related shots 300 (shown in Figure 3) by a parser 400, shown in detail in Figure 4. In the context of this invention a "shot" is a single camera "take" of a scene. A five-minute video sequence may typically have one hundred such shots or edits consisting of a series of frames F_n where, for example, $F_n = 25 \times 60 \times 5 = 7,500$ frames, whereas a ninety-minute video may have thousands of shots. If the digitised video programme is supplied with an optional edit list 202, which edit list indicates at which frames the shots 300 change, this may be utilised to divide the programme into the separate shots 300.

Basically, the parser 400 deconstructs the video into a group of sequences 321, 322, 323 (Figure 3). The sequences consist of a series of semantically related shots and, for example, one sequence may contain all the shots that feature the lead singer in a pop music video. Therefore, the function of the parser 400 is to deconstruct the programme into sequences unified by a common thread. The operation is necessary so that the tracker 800, described hereinafter, will only search for objects in sequences where they are likely to be found. The parser 400 detects shot changes, camera angle changes, wipes, dissolves and any other possible editing function or optical transition effect.

The parser 400 shown in Figure 4 receives the digital programme and the end of a shot is detected 410, e.g. by comparing edge maps of each successive frame of the video programme and stipulating that an end of shot occurs when a change in location of the edge map occurs which exceeds a predetermined threshold. The criteria 420 to be used to determine the end of a shot is input into the cut/shot detection programme by a user who is embedding data associated with an object into the video programme sequence. Information of different shots is put into an edit list 430.

A number of frames are then selected in a key-frame identifier 440 from each shot 300 to become key-frames 500 (see also Figure 5) which are representative of that shot 300. More than one key-frame may be needed for each shot where the shot 300 includes, for example, complex camera moves, such as pans or zooms, so that one key-frame 500 is not representative of the total content. Furthermore, if the video programme is of a pop group, and the sequence starts with a long shot of all the band members and speedily zooms onto the lead singer and ends with the lead singer's face filling the screen, no single frame would be representative of the whole shot, but a valid selection of three key-frames would be, for example, the first frame 311, a frame 312 about half-way through the zoom, and a final frame 313 (shown in Figure 3). Thus, key-frames 311, 312 and 313 are automatically selected which are representative of the video content of the shot 300.

As shown in Figures 3 and 4, the shots 300 are grouped into sequences by a scene grouper 450 which compares the key-frames 311 - 313 from each shot 300 with the key-frames 311 - 313 from each others shot 304, 307. This is performed by comparing the key-frames from the shots using low level features such as colour correlelograms, data maps and textures. Shots that have similar content are grouped together into a hierarchical structure by the scene grouper 450 into groups of shots having a common theme. For example, in a pop music video, it may be that there are several different sets used, but one set may appear in many places in the video. The scene grouper 450 groups sequences of

the shots 300, 304, 307 using the same set on one level and similar types of shots/sequences of the same set at another level. In this way, a hierarchical structure, termed a content tree 460, of sequences is built up. The purpose of the grouping is to aid in the selection of objects to be identified by interactive content data and also improve the efficiency of the subsequent tracking of the selected object through the video programme (described hereinafter) by ensuring that searching for a particular object is carried out only within related shots 300, 304, 307 and not through all shots of the film. The parser 400 thus assists the user to grasp the full structure and complexity of the video programme by providing a powerful browsing and object selection device as well as increasing the efficiency of the tracker by limiting tracking of an object to related shots, i.e. shots in sequences 321, 322, 323.

Having grouped the shots 300 into sequences 321, 322, 323, sequence key-frames are selected from the key-frames 311, 312, 313 of each shot to represent the sequence. A user wishing to input interactive content data representative of an object into a video programme may then use these high level key-frames to select those sequences of shots which contain objects of interest to the user. These key-frames are preferably presented to the user in a form representing the hierarchical structure in the content tree 460 of the sequences 321, 322, 323. An output 470 of the scene grouper 450 is a number of sequences of single shots, key-frame 311, 312, 313 representing the sequences and a content tree showing the hierarchical relationship between the sequences, as reflected by the key-frames.

The user intending to insert the interactive content data into the video programme views the hierarchical structure of the key frames and selects a first key-frame 311, as shown in Figure 5. In a preferred embodiment, all the key-frames may be presented to a user on a screen in miniaturised form and the user may position a cursor over the miniaturised key frame and select that key-frame. A full-sized version of the key-frame may then be presented to the user for selection of objects from the key frame 311. The user then marks with a pointing device, such as a

mouse, an object 600 within the key-frame 311 which the user intends to associate with interactive content data embedded in the programme video (as shown in Figure 6). The object may be marked by drawing a boundary box 610 around the object. To select the object 600 in the key-frame 311, the user clicks a mouse button when the cursor is at the top left corner and drags the mouse cursor to the bottom right corner of the object 600 so that the boundary box 610 is displayed around the selected object 600.

For example, to embed data information about a pop group tour date, the entire key-frame may be selected. If the key-frame contains a keyboard then the keyboard may be selected to advertise the keyboard and/or sell the keyboard on behalf of the keyboard manufacturer. Also, the lead singer who appears in the key-frame may also be selected. The boundary box shown in Figure 6 is rectangular, which is a preferred default shape, but other shapes may be used such as a parallelogram or a user defined polygon.

The selection of objects is made and the object identified 600, as shown in detail in Figure 7. Thus, the user-identifies objects 710, points to and clicks on the object 600 to provide initial object choices 715. As each object 600 is selected in the key-frame 311, attributes used to track the object through successive frames are calculated and compared with the attributes of objects already recorded 720 to ensure that the new object is distinctly different from all other objects already recorded for that frame. These attributes may include any of shape, size, position, colour, texture and intensity gradient of the object, as well as time series statistics based on these attributes. If a new object is too similar to previously recorded objects, the user is prompted for extra information about the new object. Otherwise, the attributes of the object are recorded.

The selected object in block 725 is viewed isolated from the rest of the frame. The user may then change the boundary box 610 to define the object 600 by discriminating 730 against other objects more precisely, or if two objects overlap so that they occupy the same location on the screen, the user may indicate

which object takes precedence by assigning a rank to each of the overlapping objects. For instance, in the example given above, information on the group's tour dates, which is associated with a whole frame, may be given a low rank so that, for example, any other object appearing anywhere in the frame will always have a higher rank and not be overridden by the data associated with the whole frame 311. This process is repeated for each of the key-frames 311 representing each of the sequences 321, 322, 323.

As each object is selected in the key-frame, the next step is to identify the object using data and embed the data with the object. Preferably, record addresses of data are held in a database, the data being associated with a particular object or, alternatively, instead of using a record address, the data itself may be embedded. Preferably, a graphical user interface 750 is used to drag an icon representing the data onto the object 600 within the frame 311.

Thereby the user adds the advertising content to each object in the segmented frame using a "drop and drag" technique so that, for example, an icon representing the advertiser is dragged over the object using a mouse and the relevant data is automatically embedded into the object. This process continues until all objects have been embedded with interactive data. Thereby, data representative of an object is embedded 760 into the video programme signal to provide interactive content data associated with objects 765 and a number of key-frames associated with respective embedded content data as an output 770.

Thus, the identifier 700 identifies the objects to have content embedded in them by accessing a small number of key-frames from each sequence and embedding the content.

Having embedded object descriptors in key-frames and provided content it is necessary to track the objects through the successive frames of the video programme.

Referring to Figure 8, it is necessary to track an object throughout the video programme and also as an object moves within frames and is occasionally obscured by other objects or leaves the frame being viewed, altogether. Basically, the objects are defined as a series of boundary shapes plus low-level feature

functions, e.g. shapes, edges, colour, texture and intensity gradient information. Using this representation of the objects, they are tracked through the remaining frames of the video sequence in an iterative fashion. When the plural objects have been tracked and located in every frame in which they appear, then the relevant content that was embedded in the first key-frame 311 is added automatically to the remaining frames of all sequences and this is the function of the object tracker 800, shown particularly in Figure 8.

Uncut sequences and selected objects 810 are converted 815 to a low-level representation 820 used to compare objects within a frame. For all frames, a distance measure is utilised to locate each object within each frame. A convenient distance measure is the Hausdorff measure, known per se, but this measure may be augmented with other techniques. Tracking 825 of the objects through sequential frames is iteratively provided whereby the object is initially defined in the key-frame as a two-dimensional geometric shape obtained by performing edge detection and segmenting out the edges encircled within the bounding box 610. The object 610 is then located in the next frame 312 and the attributes of the object updated to reflect the changes in position and shape that have occurred between the frames. The object with these new attributes is then located in the next frame and the process of tracker 800 continues.

Once the position of each object within all the frames of a sequence of shots has been determined, post-processing of the positions to smooth over occlusions and exits and entrances of objects is carried out.

The system is impervious to lighting changes, occlusion, camera moves, shots, breaks and optical transition effects such as wipes, fades and dissolves. The system uses a variety of known techniques to enable automatic tracking in all types of vision environments, e.g. using a group of algorithms, the selection of which is dependent upon the visual complexity of the sequence. These algorithms are known per se, although the person skilled in the art may use heuristics to optimise performance for tracking. The data added to the objects in the key-frames is

then automatically added to the object in all frames as the object is tracked throughout the entire video sequence 830.

A user may review the tracks produced and enter any corrections 835. The corrections are made by stopping the reviewed sequence at the erroneous frame, clicking on the object which is in error and dragging it to its correct position. Thus, using a graphical user interface, the video is stopped at the location in which the location of the object is incorrectly identified and the bounding box 610 is dragged and dropped at its correct location, thereby re-defining the attributes of the object for that frame and basing the definition of the object for subsequent frames on that new definition, thereby producing verified tracks 845.

Finally, all frames in all sequences of the video will have relevant objects identified and embedded with interactive content data 850.

Output from the tracker 800 is applied to a streamer 900, shown in Figure 9, in which the validity of the embedded interactive content data is checked, the order that the embedded interactive content data is output is synchronised, where necessary, with the audio/visual frames.

The streamer checks that all objects in all frames have embedded content data 850 and that the content is labelled and valid using encoder setting 920 to act upon encoder and error checker 910. Verification 940 that the content is correctly labelled and valid occurs and the output 930 may be either a complete broadcasting compliant transport stream, such as MPEG-2/DVB audio, video and embedded objects and content data, or as embedded objects and content data alone.

The streamer 900 must determine in which of three categories the embedded content data falls, namely frame-synchronous data, segment-synchronous data, or data to be streamed just-in-time. Frame synchronous data consists of the object positions, shapes, ranks and pointers to a table of pointers to data may be associated with the correct frame number in the video programme from source 201. Segment-synchronous data is used to update the table of pointers to embedded content data so that when objects

change, the embedded data changes. This data may be associated with the frame number at which the content changes. Data to be streamed "just in time" must be streamed to the end user before it is required by any of the objects. This transport stream is then packetised into MPEG-2/DVB compliant packets.

If a fully embedded audio visual programme is required, the packetised transport stream and the video programme are multiplexed together, as shown in Figure 10.

Referring to Figure 10, the different elements that constitute the embedded video programme are combined into a single transport stream 1001 in preparation for broadcasting by a network operator. The programme consists of a video stream 1010, an audio stream 1020, both of which streams are uncompressed. Both the video data 1010 and the audio data 1020 are encoded and compressed in respective MPEG-2 elementary encoders 1015 and 1025 to produce elementary streams of data 1030, 1035 respectively. MPEG-2 compliant data sequence 930 is error checked 1037 to produce an elementary stream of data 1040. The elementary streams 1030, 1035 and 1040 are applied to packetisers 1050, 1055 and 1060, which each accumulate data into fixed size blocks to which is added a protocol header. The output from the packetisers is termed a packetised elementary stream (PES) 1070. The packetised elementary streams 1070, in combination with digital control data (PSI) 1075, is applied to a systems layer multiplexer 1080 having a systems clock 1085. The PES packet is a mechanism to convert continuous elementary streams of information 1030, 1035 and data sequence 930 into a stream of packets. Once embedded in PES packets the elementary streams may be synchronised with time stamps. This is necessary to enable the receiver (PC or TV) to determine the relationship between all the video, audio and data streams that constitute the embedded video programme.

Each PES packet is fed to the system multiplexer 1080. There the packets are encapsulated into transport packets to form the transport stream 1001 that is used for broadcast. In this respect, the transport stream 1001 carries packets in 188 byte blocks and the transport stream 1001 constitutes a full so-called

eMUSE channel that is fed to the network operator for broadcast. In essence, the transport stream is a general purpose way of combining multiple streams using fixed length packets.

The structure of a packet is shown in Figure 11. The packet 1100 shown in Figure 11 has a header 1110 with a synchronisation byte, a 13-bit packet identifier (PID) and a set of flags to indicate how the packet should be processed. The transport multiplexer assigns a different packet identifier to each PES 1070 to uniquely identify the individual streams. In this way, the packetised data sequence 930 is uniquely identified. The synchronisation of the elementary streams is facilitated by sending time stamps in the transport stream 1001.

Two types of time stamps may be used:

1. A reference time stamp to indicate the current time, that is clock 1085 information, and
2. A decoding time stamp.

The decoding time stamps are inserted into the PES to indicate the exact time when the data stream has to be synchronised with the video and audio streams. The decoding time stamp relies on the reference time stamp for operation. After the transport stream has been broadcast, the PC or TV uses the time stamps to process the data sequence in relation to the video and audio streams.

In order for the receiver (PC or TV) to know how to decode the channel, it needs to access a set of signalling tables known as Programme Specific Information (PSI) labels which are sent as separate packets within the transport stream 1001 with their own PID tables. There are two tables that are needed to enable the receiver to decode a channel. The first is the programme association table (PAT) 1130 which lists all the channels that are available within the satellite broadcast and has a packet ID (PID) value of 0 which makes it easy to identify. In the example, the eMUSE channel, i.e. the channel carrying the video programme, is represented as PID 111.

A programme table map (PMT) 1140 identifies all the elementary streams contained in the embedded video signal. Each elementary stream is identified by a PID value, e.g. video from

video camera 1 is PID 71. The data sequence 930 has a PID value 92 in the example of Figure 11. The receiver video and audio decoders search the PMT table to find the appropriate packets to decode. Similarly, the programme for retrieving the embedded data searches the PMT to find the data sequence which, in the example of Figure 11, is PID 92. The data retrieval programme then filters out these packets and synchronises them with the appropriate video and audio to enable the user to select the various objects.

Having embedded the interactive content data into the video programme signal, it is broadcast and the manner of reception and retrieval of the data will now be explained with reference to Figure 12.

Hardware is provided on a satellite receiver card 1210 which resides on the user's PC 40 or digital set top box 50 and software allows the viewer to interact with the dynamic objects in the broadcast, for example to facilitate Internet access and Internet browsers, such as Internet Explorer and Netscape and, for TV applications, is compatible with Sun's Open TV operating system.

The received MPEG-2/DVB signal is separated into MPEG-2 video 1215, MPEG-2 audio 1220 and the data sequence 930 and the decoded video 1225, audio and data sequence is applied to a synchroniser 1230. Output from the synchroniser comprising the video programme with embedded interactive content data is displayed 1240 by the PC VDU or TV screen.

A user clicks a mouse 56 or presses a remote control button at a frame containing an object of interest, which causes the display on the screen to split in two. For example, on the left hand screen, the video programme continues to run as normal and, on the right hand screen, the objects present in the frame which was active the time the mouse was clicked, are displayed as cut-outs, with the intervening spaces blanked out. The user then clicks on the object of interest to see which advertisers it represents, e.g. if the user clicks on the lead singer, then the screen will display the lead singer only and a textual list of advertisers or an icon-based display of advertisers will be

viewed. If the user clicks on the advertiser's name or icon, the user goes directly to view the advertised products.

After interacting with the site the user may decide to purchase the product via an e-commerce transaction. Further, if the user clicks on the suit of the lead singer, the entire catalogue of the suit manufacturer may be made available as part of the streamed digital broadcast. This return path via the Internet is purely to facilitate a transaction as the data sequence 930 initiates the push technology approach to streaming advertising information once the user has selected amongst the numerous objects within the frame.

Although the user can interact with the broadcast in such an on-line manner as described above, alternatively, the data may be viewed off-line, i.e. while a viewer continues to watch a programme, the user may select various frames during the broadcast and store the frames for later retrieval of the associated data. Where there is not sufficient local memory to store the data, addresses of the data in local or remote databases, e.g. Web sites, are stored and the end user is able to subsequently access the databases to retrieve the data. The user then selects with the mouse or the remote control the object 600 of interest and another screen may then be displayed showing the object 600 and a menu of data elements associated with that object. The user clicks one of the menu items and is able to directly view data on the advertised product or be given access to a Web site over the Internet. Alternatively, as soon as a user selects a menu item, a catalogue may be viewed which has been embedded in the broadcast signal.

The data which the end user accesses may be streamed with a broadcast signal or may be held in a local data base which may be pre-loaded into the end user's device prior to viewing the video sequence. When viewing information streamed with a broadcast, the information associated with a particular programme is streamed in parallel with the programme and stored locally. When the user selects an object, this local data is viewed.

Claims:

1. An interactive system including means for providing a video programme signal, means for generating interactive content data associated with at least one object, said data being associated with frames of said video programme signal in which the object appears, means for multiplexing said data with said video programme signal, means for viewing the video programme signal, means for retrieving said data and means for using said data to obtain details of said object.
2. An interactive system claimed in claim 1, wherein each frame of said video programme includes said interactive content data.
3. An interactive system as claimed in claims 1 or 2, wherein said means for using said data further include means for producing a list of details of said object and means for selecting from said list.
4. An interactive system as claimed in any of claims 1 to 3, wherein said means for using said data include means for accessing an interactive Web site to obtain said details of said object.
5. An interactive system as claimed in claims 3 or 4, wherein said means for accessing an interactive Web site is adapted to secure details of said object which may include a purchasing transaction for said object or browsing an advertising catalogue.
6. An interactive system as claimed in any of the preceding claims, wherein the means for generating includes means for tracking said object in each frame of said video programme signal in which said object appears and means for identifying the location of said object in each said frame.
7. An interactive system as claimed in claim 6, wherein said tracking means includes means for determining scene breaks and

means for searching for said object in a next frame in which said object appears.

8. An interactive system as claimed in any of the preceding
5 claims, wherein said multiplexing means includes means for synchronising said data with audio and video data of said programme signal to generate a transport stream.

9. An interactive system as claimed in claim 8, wherein said
10 system includes means for broadcasting said transport stream via, at least one of a satellite, terrestrial and cable network.

10. An interactive system as claimed in any of the preceding
15 claims, wherein said means for retrieving includes one of a mouse, a keyboard, and remote control device.

11. An apparatus for associating data representative of an object with a digital video programme including means for providing a digital video programme having plural individual
20 frames at least some of which incorporate said object, means for selecting a frame of the video programme in which said object appears to provide a key-frame, means for selecting said object within the key-frame with which data is to be associated, means for extracting attributes of the object from the key-frame, means
25 for associating interactive data with the object in the key-frame, means for utilising the attributes of the object for tracking the object through subsequent frames of the video programme, whereby said interactive data is associated with the object in subsequent frames of the video programme in which said
30 object has been tracked and said interactive content data is embedded with data representative of said object in a data sequence.

12. An apparatus as claimed in claim 11, wherein means are
35 provided for converting said data sequence to a standard data sequence.

13. An apparatus as claimed in claims 11 or 12, including means for converting a video programme in an analogue format to digitised form.

5 14. An apparatus as claimed in any of claims 11 to 13, wherein the means for selecting a frame of the video programme includes means for producing an edit list to divide the digitised video programme into a plurality of sequences of related shots, and means for selecting at least one key-frame from within each
10 sequence.

15. An apparatus as claimed in claims 14, wherein the means for producing an edit list further includes means for parsing the video programme by identifying separate shots in the video
15 programme to produce the edit list, means for identifying shots containing related content to form a sequence of shots containing related content, and means for producing a hierarchy of groups of shots.

20 16. An apparatus as claimed in claim 15, wherein said means for parsing include means for inputting criteria to be used to recognise a change of shot.

25 17. An apparatus as claimed in any of claims 11 to 16, wherein the means for extracting attributes of the object includes means for isolating the object within a boundary formed on the frame, means for performing edge detection within the boundary to identify and locate edges of said object, and storing means for storing a geometric model of said object.

30 18. An apparatus as claimed in any of claims 11 to 17, wherein said means for extracting attributes of said object also includes means for recording at least one of the attributes of shape, size, position, colour, texture, intensity gradient of said
35 object, and time series statistics based on said attributes.

19. An apparatus as claimed in any of claims 11 to 18, wherein said means for extracting attributes of said object includes means for comparing said attributes of said object with attributes of objects previously stored to determine whether the object is distinguishable therefrom, and when said object is determined not to be distinguishable, providing means for re-defining the object.

20. An apparatus as claimed in any of claims 11 to 19, wherein said means for extracting said attributes includes means for comparing the location in the frame of said object with the location of objects already stored for that frame to determine whether that object is distinguishable therefrom, and where the location of said object is not distinguishable from the location of another object providing means for assigning rank to the objects to determine which object will be associated with that location.

21. An apparatus, as claimed in any of claims 11 to 20, wherein the means for tracking the object includes means for updating the stored attributes of the object as the object moves location within different frames.

22. An apparatus as claimed in any of claims 11 to 21, wherein said means for tracking includes plural algorithm means for use depending on the visual complexity of a sequence to automatically track objects in different types of visual environment.

23. An apparatus as claimed in any of claims 11 to 22, wherein said tracking means includes means for converting all the frames to be tracked to a low-level representation, means for determining the position of each object in the frames by minimising a distance measure to locate each object in each frame, means for processing the positions of said object to smooth over occlusions and the entrances and exits of objects into and out of said frames, and means for reviewing the object

within a tracked sequence and for correcting the location attributes of any misplaced objects.

24. An apparatus, as claimed in any of claims 11 to 23, wherein the means for associating includes means for providing a database of different types of data including one or more of URLs, HTML pages, video clips, audio clips, text files and multimedia catalogues, and means for selecting said interactive content data from the database to associate with said object.

25. An apparatus, as claimed in any of claims 11 to 24, wherein the means for associating produces said data sequence using means for determining whether the embedded interactive content data is frame synchronous data associated with object positions, shapes, ranks and pointers in a frame, or group-synchronous data associated with all the objects in a group, or is data to be streamed just in time, wherein means are provided for associating frame synchronous data with the corresponding frame, means are provided for associating group synchronous data with the frame at which a group changes, and means are provided for streaming just in time data to a user before it is required to be associated with the corresponding objects.

26. An apparatus as claimed in any of claims 11 to 25, wherein means are provided to associate different interactive content data with respectively different objects.

27. An apparatus for embedding a data sequence within a generic digital transport stream, including means for receiving a data sequence of interactive content data associated with an object in a digitised video signal, means for synchronising the data sequence with the video and audio of the digitised video signal to generate a further transport stream, and means for associating a packet identifier with the further transport stream.

28. An apparatus as claimed in claim 27, wherein means are provided for broadcasting the further transport stream to viewers.

5 29. An apparatus as claimed in claims 27 or 28, wherein the means for receiving a data sequence includes means for receiving elementary streams comprising a digital video signal stream, a digital audio stream, a digital data sequence stream and a digital control data stream, means for packetising each of the
10 data streams into fixed size blocks and adding a protocol header to produce packetised elementary streams, and means for synchronising the packetised elementary streams with time stamps to establish a relationship between the data streams.

15 30. An apparatus as claimed in any of claims 27 to 29, wherein the means for synchronising the data sequence includes means for multiplexing packetised elementary streams into transport packets headed by a synchronisation byte, and means for assigning a different packet identifier to each packetised elementary stream.

20 31. An apparatus as claimed in claim 30, wherein means for synchronising the packetised elementary streams with time stamps includes means for stamping with a reference time stamp to indicate current time, and means for stamping with a decoding
25 time stamp to indicate when the data sequence stream has to be synchronised with the video and audio streams.

32. An apparatus as claimed in claim 28, wherein the means for broadcasting the further transport streams to users includes
30 means for providing a programme association table listing all the channels to be available in the broadcast, means for providing a programme map table identifying all the elementary streams in the broadcast channel, and means for transmitting the programme association table and the programme map table as separate packets
35 within the further transport stream.

33. An apparatus for retrieving data embedded in a generic digital transport stream in which the embedded data includes a data sequence of data associated with objects represented by the generic digital transport stream, said apparatus including means
5 for recognising a packet identifier within the video signal, means for extracting the data sequence from the generic digital transport stream, means for identifying objects within the video sequence from which to retrieve associated data, means for
10 synchronising said data sequence to said identified objects and means for interactively using said associated data.

34. An apparatus as claimed in claim 33, wherein said means for identifying objects includes means for selecting an object within a frame, means for displaying data associated with said object,
15 means for selecting data from a list of displayed data, and means for extracting the embedded data associated with the data relating to said object.

35. An apparatus as claimed in claims 33 or 34, wherein means are
20 provided for selecting a frame to display the objects having embedded associated data, means for selecting one of the displayed objects to display a list of the data associated with said object, and means for selecting from said list.

25 36. An apparatus as claimed in claim 35, wherein the means for selecting a frame includes means for storing the frame for subsequent display and subsequent recall of the frame.

30 37. An apparatus as claimed in any of claims 33 to 36, wherein the extracted embedded data is applied to means for accessing an Internet web site to facilitate interactive communication such as e-commerce.

1 / 9

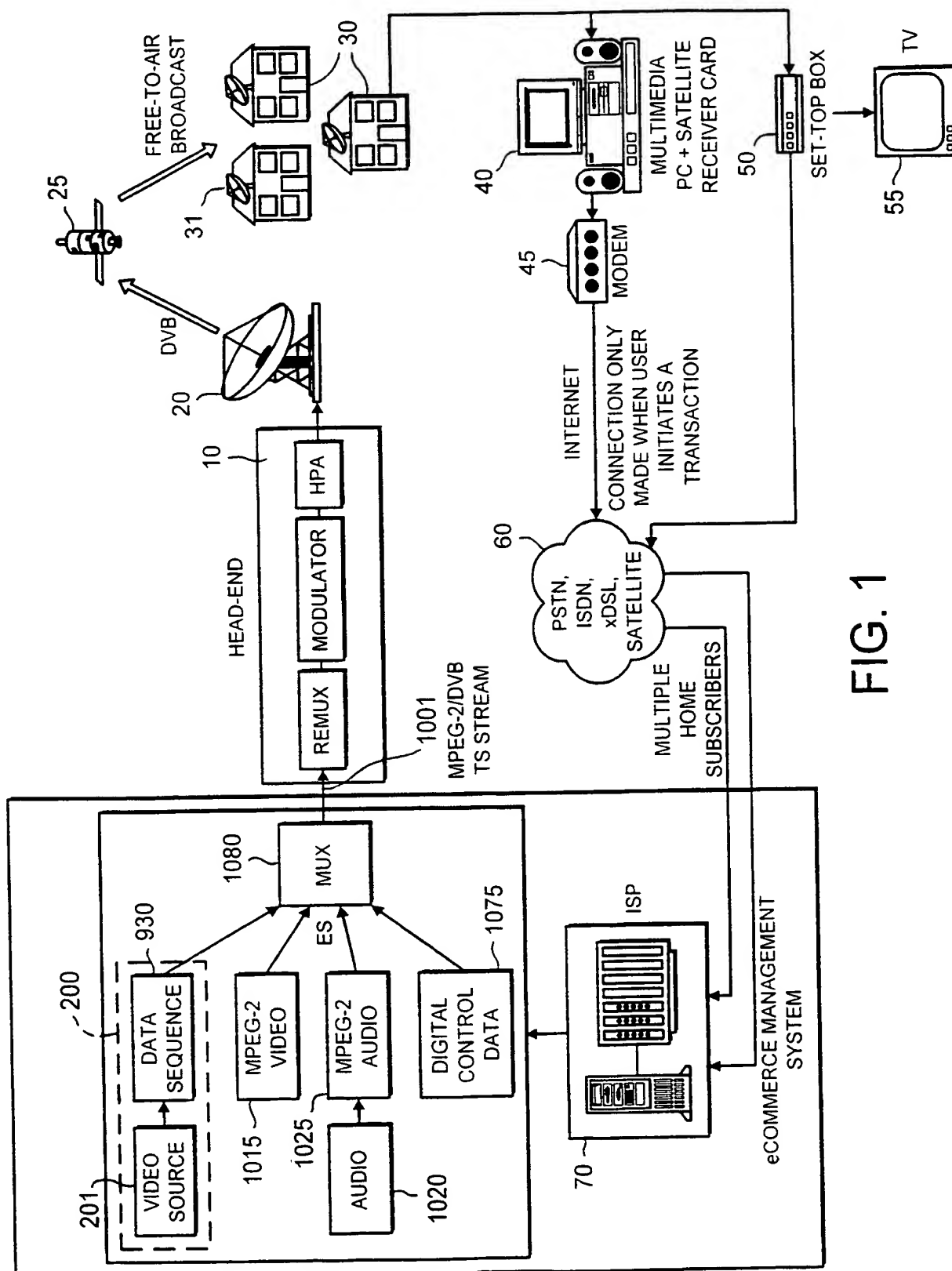


FIG. 1

2 / 9

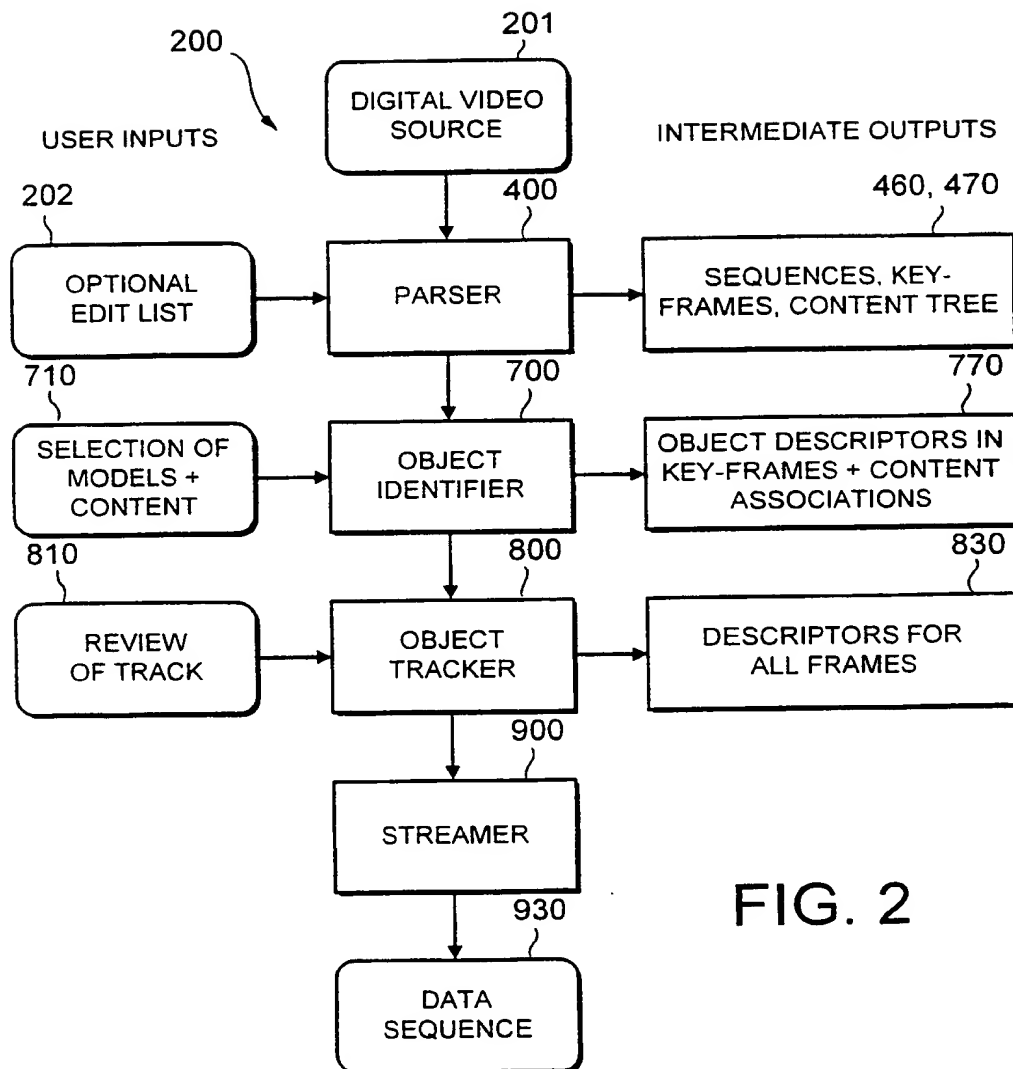


FIG. 2

3 / 9

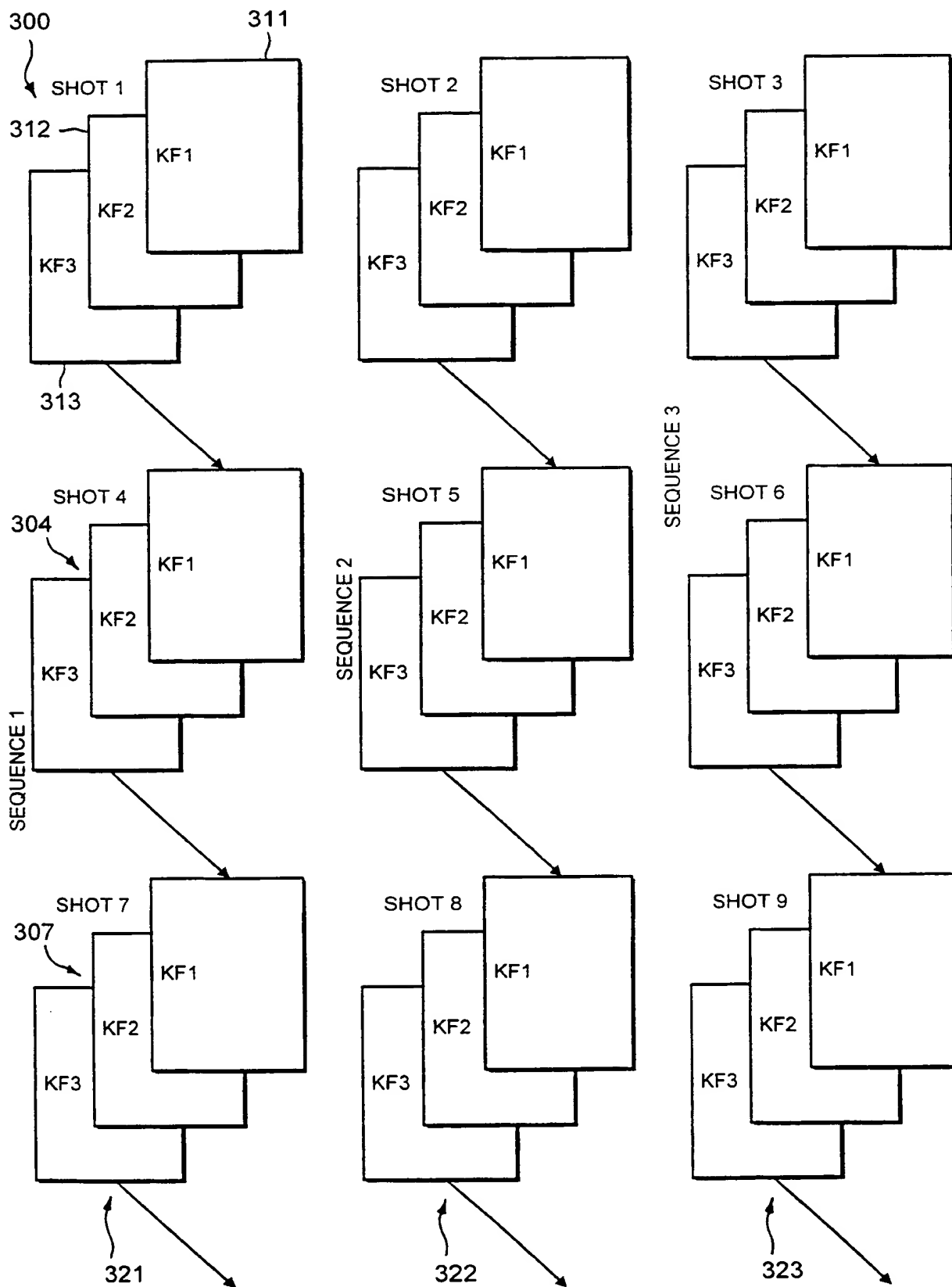
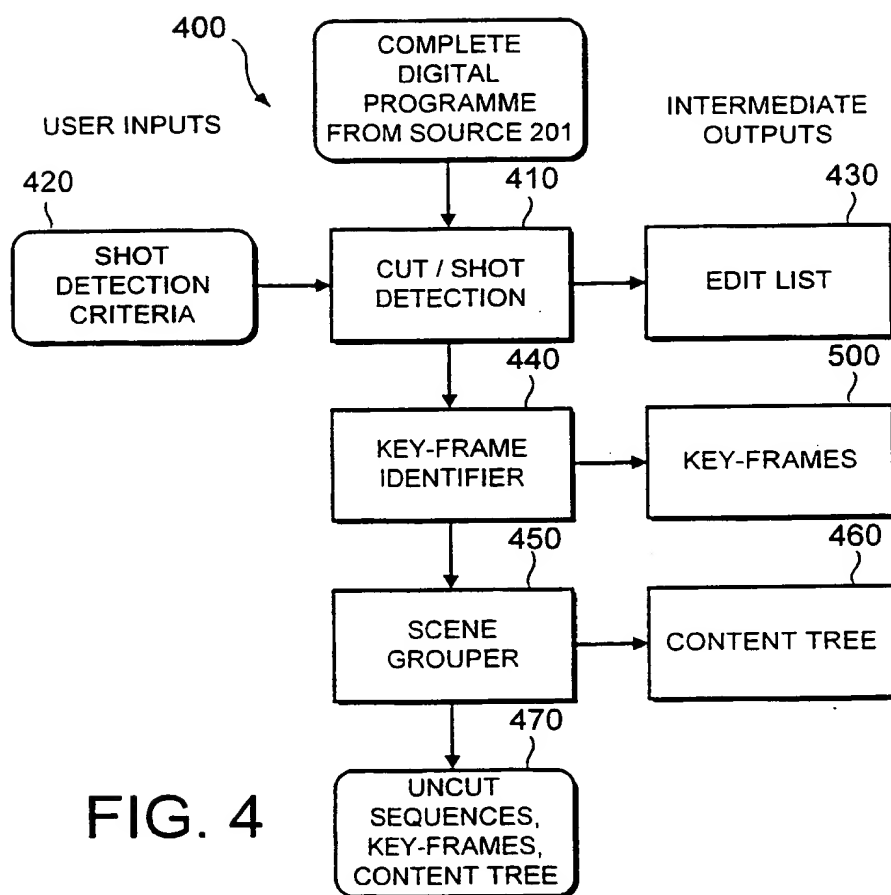


FIG. 3

4 / 9



5 / 9

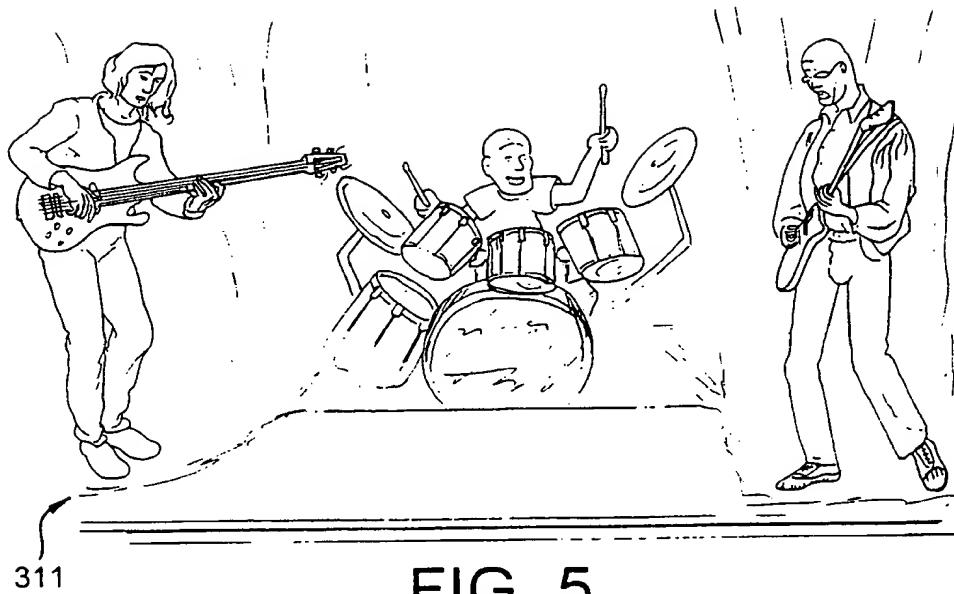


FIG. 5

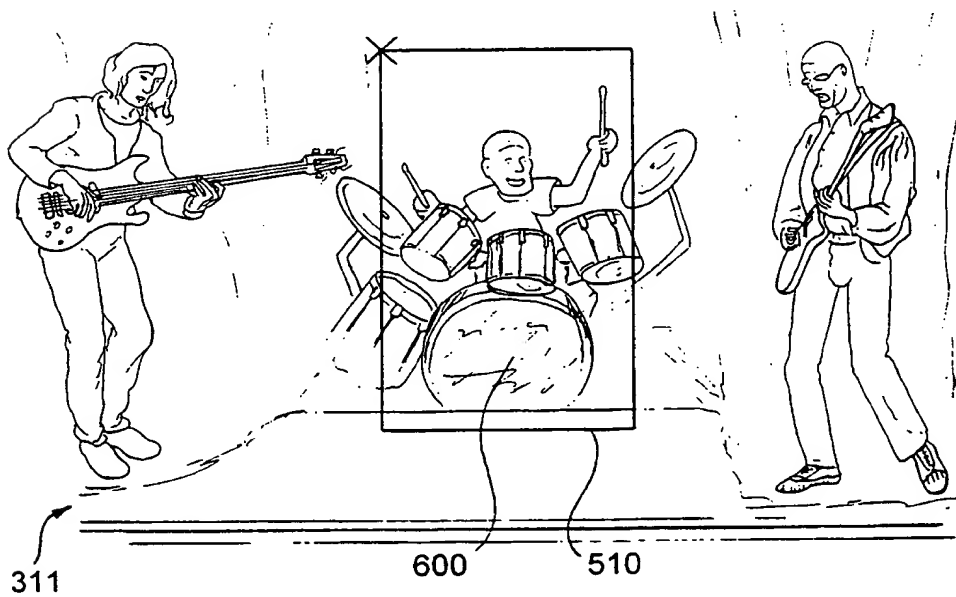


FIG. 6

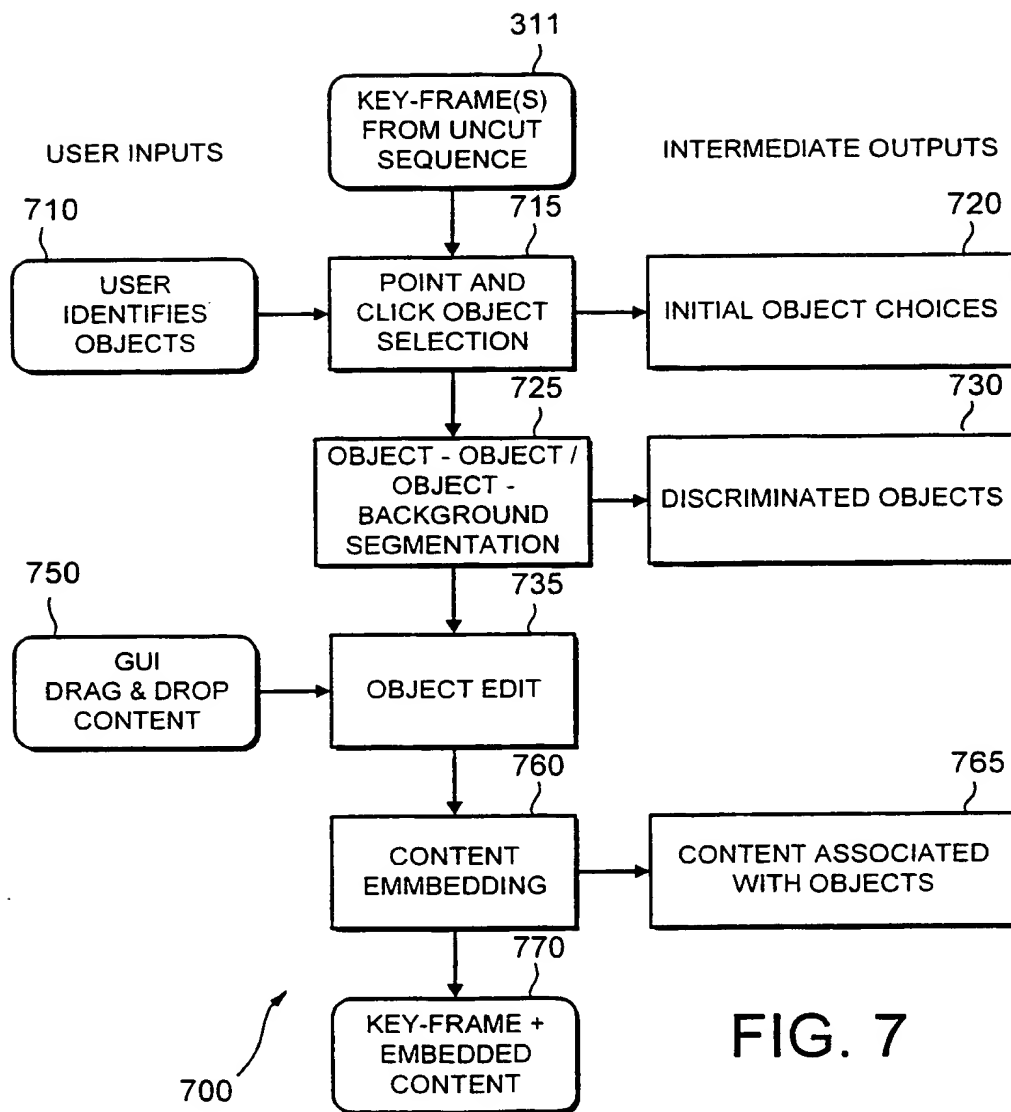


FIG. 7

7 / 9

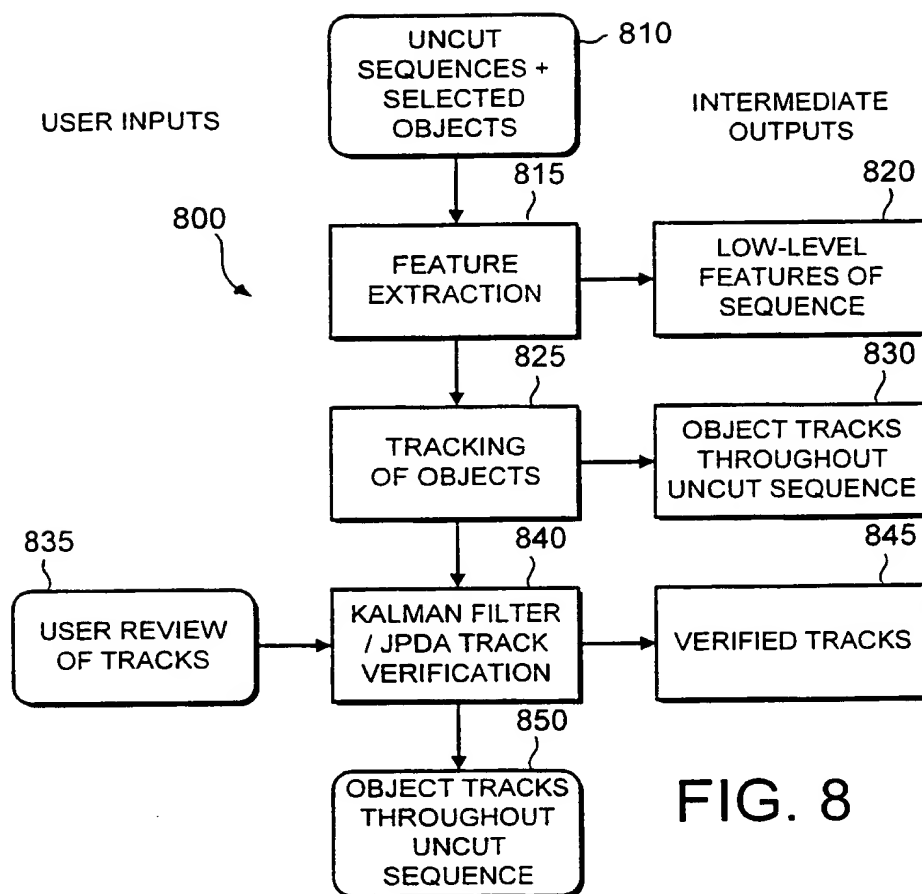


FIG. 8

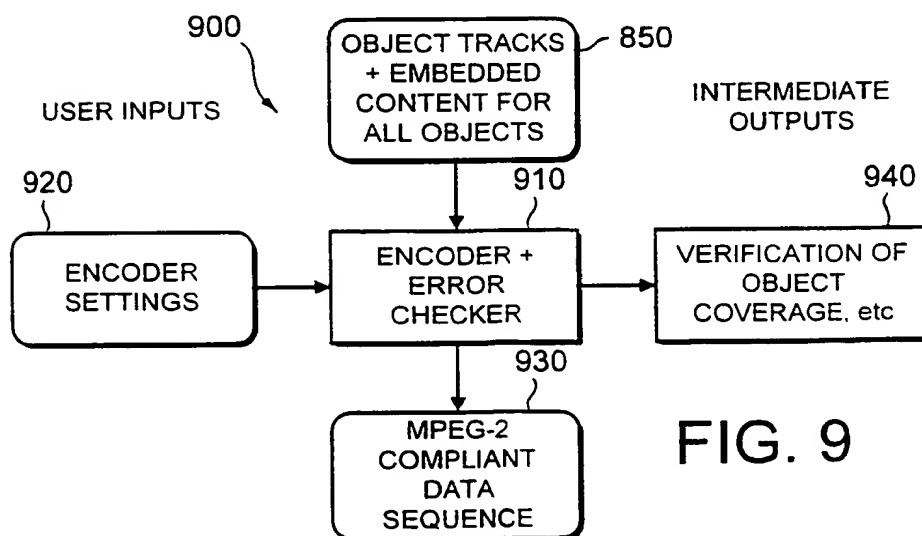


FIG. 9

8 / 9

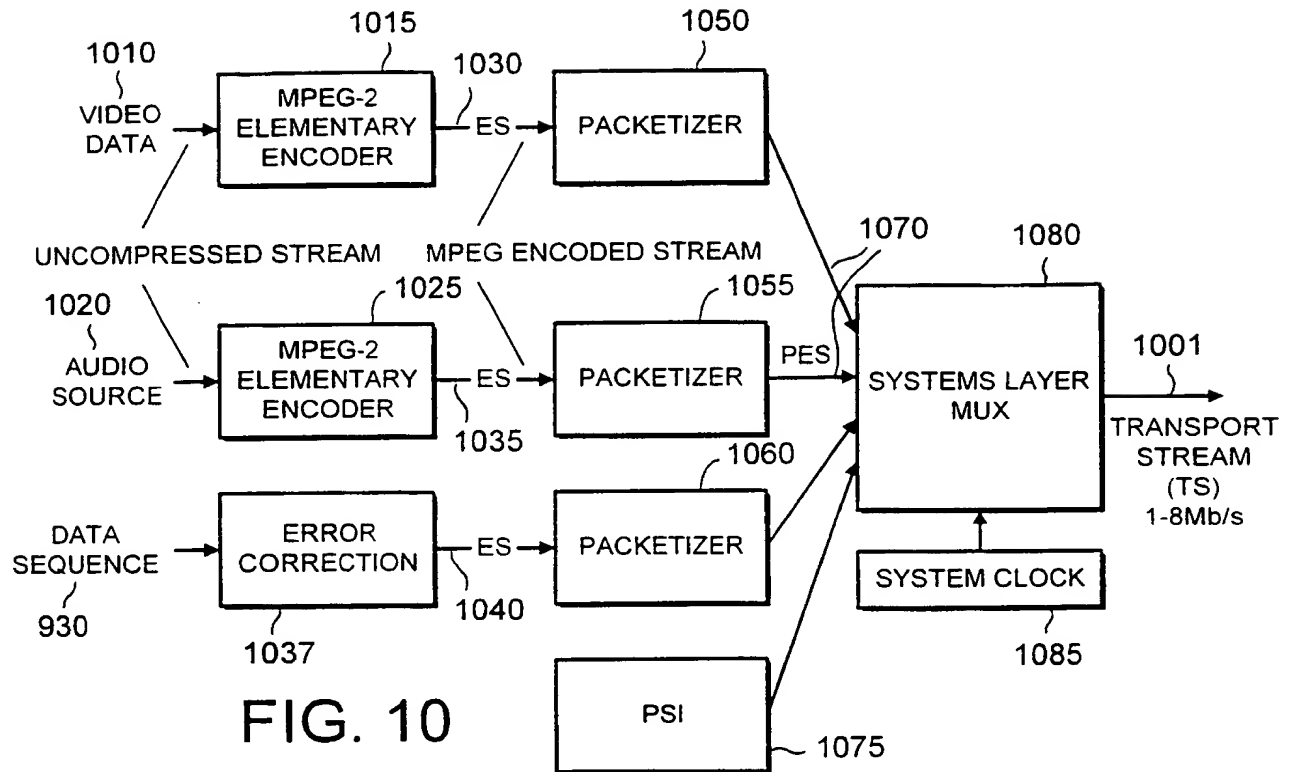


FIG. 10

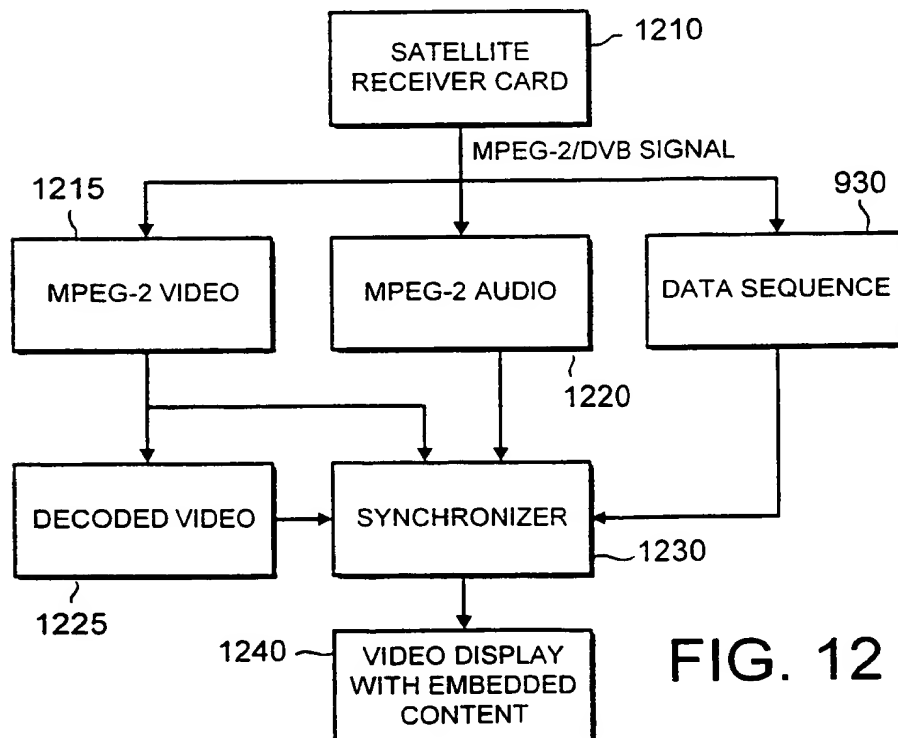


FIG. 12

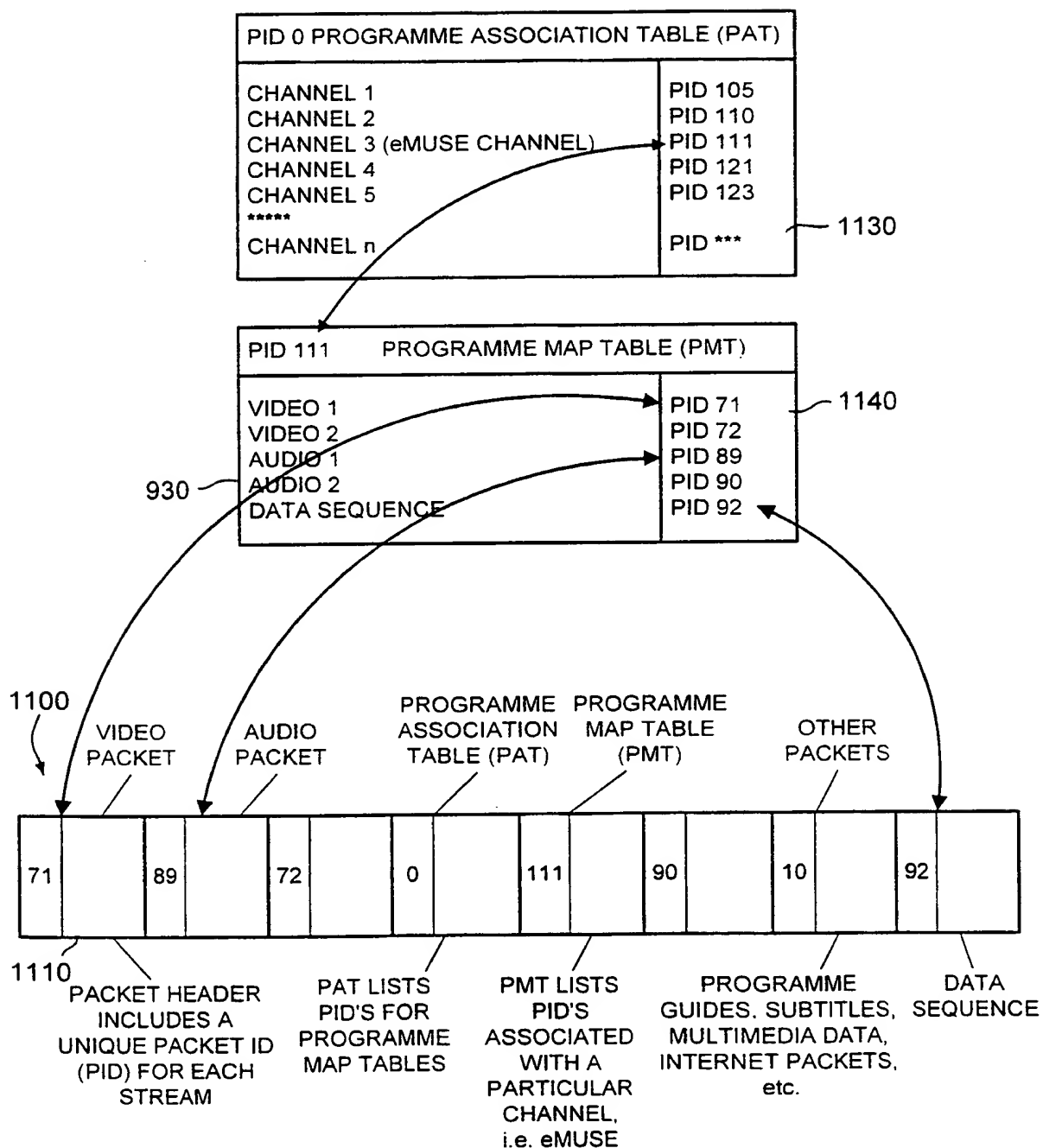


FIG. 11

(19) World Intellectual Property Organization
International Bureau



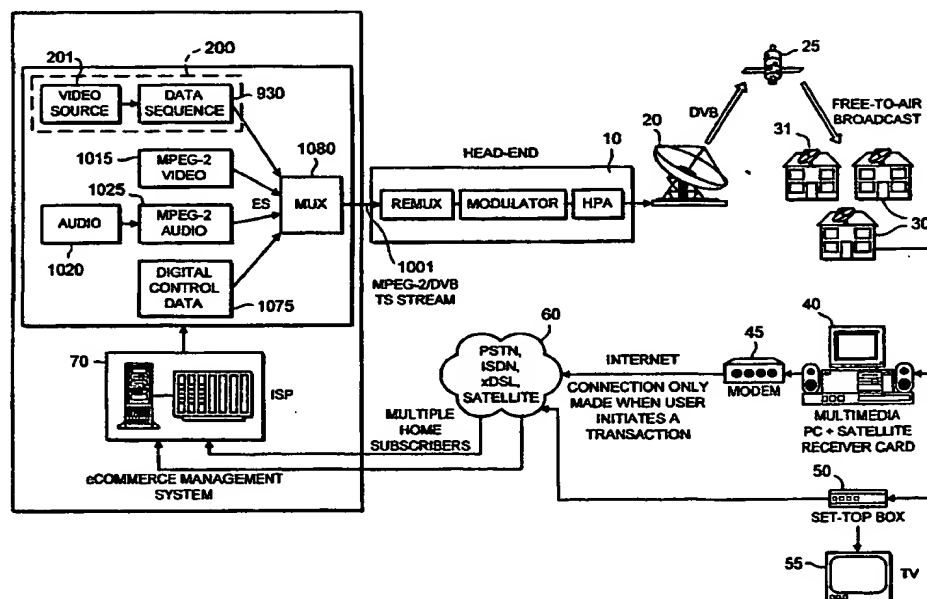
(43) International Publication Date
3 August 2000 (03.08.2000)

PCT

(10) International Publication Number
WO 00/45599 A3

- (51) International Patent Classification⁷: **H04N 7/24**
- (21) International Application Number: **PCT/IB00/00135**
- (22) International Filing Date: **28 January 2000 (28.01.2000)**
- (25) Filing Language: **English**
- (26) Publication Language: **English**
- (30) Priority Data:
9902235.2 1 February 1999 (01.02.1999) GB
- (71) Applicant (for all designated States except US): **EMUSE CORPORATION [IE/IE]; 80 Park Avenue, Sandymount, Dublin 4 (IE).**
- (72) Inventor; and
- (75) Inventor/Applicant (for US only): **RAINSFORD, Patrick, J. [IE/IE]; 80 Park Avenue, Sandymount, Dublin 4 (IE).**
- (74) Agents: **BAILLIE, IAIN, C. et al.; Langner Parry, 52-54 High Holborn, London WC1V 6RR (GB).**
- (81) Designated States (national): **AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.**
- (84) Designated States (regional): **ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).**
- Published:
— With international search report.
- (88) Date of publication of the international search report:
28 December 2000
- For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: **INTERACTIVE SYSTEM**



(57) Abstract: An interactive system provides a video programme signal and generates interactive content data to be associated with at least one object within a frame of the video programme. The interactive content data is embedded with the object and the object is tracked through a sequence of frames and the interactive content data is embedded into each one of the frames. The programme frames with the embedded data are multiplexed with video and audio signals and may be broadcast. A receiver identifies an object of interest and the embedded data associated with the object is retrieved. The embedded data may be used for e-commerce.

INTERNATIONAL SEARCH REPORT

International Application No

PCT/IB 00/00135

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 H04N/24

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHEDMinimum documentation searched (classification system followed by classification symbols)
IPC 7 H04N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

WPI Data, PAJ, INSPEC

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 97 41690 A (AWARD SOFTWARE INTERNATIONAL I) 6 November 1997 (1997-11-06)	1,2,4-6, 10
Y	page 2, line 26 -page 7, line 28	3,7
A	page 9, line 20 -page 16, line 18 abstract; figures 2A,2B,2C,3A,3B	8,9, 11-37
Y	EP 0 851 681 A (HITACHI LTD) 1 July 1998 (1998-07-01)	3,34-36
A	column 2, line 44 -column 7, line 14 column 9, line 5 -column 19, line 54 abstract; figures 1,7A,7B,7C,10	1,2,4,5, 8-10, 27-33

	-/--	

☒ Further documents are listed in the continuation of box C.☒ Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search

7 August 2000

Date of mailing of the international search report

11.09.2000

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Authorized officer

La, V

INTERNATIONAL SEARCH REPORT

International Application No.

PCT/IB 00/00135

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	POPE A ET AL: "Video abstraction: summarizing video content for retrieval and visualization" CONFERENCE RECORD OF THIRTY-SECOND ASILOMAR CONFERENCE ON SIGNALS, SYSTEMS AND COMPUTERS (CAT. NO.98CH36284), CONFERENCE RECORD OF THIRTY-SECOND ASILOMAR CONFERENCE ON SIGNALS, SYSTEMS AND COMPUTERS, PACIFIC GROVE, CA, USA, 1-4 NOV. 1998, pages 915-919 vol.1, XP002139169 1998, Piscataway, NJ, USA, IEEE, USA ISBN: 0-7803-5148-7 section 1."Introduction" section 2.1 "Generic schema" page 917, left-hand column, line 27 -right-hand column, line 9 page 918, left-hand column, line 6 - line 19	7,14-16, 23
A	abstract; figure 1 ---	1,6,11
X	US 5 708 845 A (WISTENDAHL DOUGLASS A ET AL) 13 January 1998 (1998-01-13)	11-13, 17-19, 21,22, 24-26
Y	the whole document	14-16, 20,23
Y	EP 0 675 461 A (CASIO COMPUTER CO LTD) 4 October 1995 (1995-10-04) column 21, line 15 - line 29	20
A	abstract ---	11
X	WO 98 57499 A (SHARP KK) 17 December 1998 (1998-12-17)	1,8,9, 27,28, 30,32, 33,37
Y	page 6, line 2 -page 7, line 4 page 7, line 27 -page 12, line 16 abstract --- -/--	29,31, 34-36

INTERNATIONAL SEARCH REPORT

(International Application No

PCT/IB 00/00135

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	SARGINSON P A: "MPEG-2: A TUTORIAL INTRODUCTION TO THE SYSTEMS LAYER" IEE COLLOQUIUM ON MPEG WHAT IT IS AND WHAT IT ISN'T, GB, IEE, LONDON, 1995, pages 4-1-4-13, XP000560804 section "The Basics" section "Packetised Elementary Streams" sections "The Transport Stream multiplex", "The transport packet header", "Programme Specific Information"	29,31
A	section "Time stamps and clock references : the basics", "Time stamps and clock references : the details"	1,8,9, 27,28, 30,32
X	WO 98 43437 A (DECLERCK CHRISTOPHE ;CANAL PLUS SA (FR); CHAUMET CHRISTOPHE (FR)) 1 October 1998 (1998-10-01) page 1, line 1 -page 8, line 3 page 12, line 22 -page 18, line 20 abstract	1,10
A		2-5,8,9, 27-32
X	US 5 818 935 A (MAA CHIA-YIU) 6 October 1998 (1998-10-06) column 2, line 48 -column 3, line 59 column 6, line 7 - line 39 column 14, line 61 -column 15, line 67 column 17, line 23 -column 18, line 20 column 19, line 11 -column 20, line 53 abstract	1,4,5,10
A		2,3,8,9, 27-32
X	"MULTIMEDIA HYPERVIDEO LINKS FOR FULL MOTION VIDEOS" IBM TECHNICAL DISCLOSURE BULLETIN,US,IBM CORP. NEW YORK, vol. 37, no. 4A, 1 April 1994 (1994-04-01), page 95 XP000446196 ISSN: 0018-8689 the whole document	11,18
A	EP 0 596 823 A (IBM) 11 May 1994 (1994-05-11) the whole document	1-37
A	WO 98 00976 A (SUN MICROSYSTEMS INC ;THOMSON CONSUMER ELECTRONICS (US)) 8 January 1998 (1998-01-08) page 3, line 15 - line 27 abstract; figures 5,6	1-7,10, 33-37

-/--

INTERNATIONAL SEARCH REPORT

International Application No

PCT/IB 00/00135

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 97 19560 A (IMEDIA CORP) 29 May 1997 (1997-05-29) page 8, line 3 -page 11, line 29 page 14, line 3 -page 24, line 27 abstract; figures 1,4,6	1-10, 27-37
A	OLSON C F ET AL: "AUTOMATIC TARGET RECOGNITION BY MATCHING ORIENTED EDGE PIXELS" IEEE TRANSACTIONS ON IMAGE PROCESSING,US,IEEE INC. NEW YORK, vol. 6, no. 1, 1997, pages 103-113, XP000642455 ISSN: 1057-7149 section II "Matching oriented edge pixels" abstract	11-26
A	GUNSEL B ET AL: "TEMPORAL VIDEO SEGMENTATION USING UNSUPERVISED CLUSTERING AND SEMANTIC OBJECT TRACKING" JOURNAL OF ELECTRONIC IMAGING,US,SPIE + IS&T, vol. 7, no. 3, 1 July 1998 (1998-07-01), pages 592-604, XP000771766 ISSN: 1017-9909 the whole document	11-26
A	WO 98 21688 A (SARNOFF CORP) 22 May 1998 (1998-05-22) page 16, line 29 -page 20, line 10 abstract	11-26

INTERNATIONAL SEARCH REPORT

International application No.
PCT/IB 00/00135

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
2. ☐ Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:
3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. ☒ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.
- ☒ No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. Claims: 1-7,10

An interactive system including a means for providing a video programme signal, a means for generating interactive content data associated with at least one object, said data being associated with frames of said video programme signal in which the object appears, a means for multiplexing said data with the video program signal, a means for viewing the video programme signal, a means for retrieving said data and a means for using said data to obtains details of the object, said using means includes a means for producing a list of details of said object and a means for selecting from said list.

2. Claims: 8,9,27-32

An apparatus for embedding a data sequence within a generic digital transport stream, including a means for receiving a data sequence of interactive content data associated with an object in a digitised video signal, a means for synchronising the data sequence with the video and audio of the digitised video signal to generate a further transport stream, and a means for associating a packet identifier with the further transport stream, wherein the means for receiving a data sequence includes a means for receiving elementary streams comprising a digital video signal stream, a digital audio stream, a digital data sequence and a digital control data stream, a means for packetising each of the data streams into fixed sized blocks and adding a protocol header to produce packetised elementary streams, and means for synchronising the packetised elementary streams with time stamps to establish a relationship between the data streams.

3. Claims: 11-26

An apparatus for associating data representative of an object with a digital video programme including a means for providing a digital video programme having plural individual frames at least some of which incorporate said object, a means for selecting a frame of the video programme in which said object appears to provide a key-frame, a means for selecting said object within the key-frame with which data is to be associated, a means for extracting attributes of the object from the key-frame, a means for associating interactive data with the object in the key-frame, a means for utilising the attributes of the object for tracking the object through subsequent frames of the video programme, whereby said interactive data are associated with the object in subsequent frames in which the object has been tracked

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

and said interactive data are embedded with data representative of said object in a data sequence.

4. Claims: 33-37

An apparatus for retrieving data embedded in a generic digital transport stream in which the embedded data includes a data sequence of data associated with objects of the generic digital transport stream, comprising a means for recognizing a packet identifier within the transport stream, a means for extracting the data sequence from the transport stream, a means for identifying objects within the video sequence from which to retrieve associated data, a means for synchronising said data sequence to said identified objects and a means for interactively using said associated data, wherein a means is provided for selecting a frame to display the objects having embedded associated data and for selecting one of the displayed objects.

Information on patent family members

PCT/IB 00/00135

Form PCT/ISA/210 (patent family annex) (July 1992)

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/IB 00/00135

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9843437 A		EP 0968611 A	05-01-2000
		EP 0968608 A	05-01-2000
		EP 1010068 A	21-06-2000
		EP 0968469 A	05-01-2000
		EP 0972406 A	19-01-2000
		NO 994529 A	19-11-1999
		NO 994530 A	19-11-1999
		NO 994531 A	19-11-1999
		NO 994532 A	19-11-1999
		NO 994533 A	22-11-1999
		NO 994535 A	22-11-1999
		NO 994536 A	22-11-1999
US 5818935 A	06-10-1998	NONE	
EP 0596823 A	11-05-1994	US 5539871 A	23-07-1996
		JP 2677754 B	17-11-1997
		JP 7085243 A	31-03-1995
WO 9800976 A	08-01-1998	US 5903816 A	11-05-1999
		AU 3648797 A	21-01-1998
		EP 0909513 A	21-04-1999
WO 9719560 A	29-05-1997	US 5966120 A	12-10-1999
		AU 7724996 A	11-06-1997
		EP 0862836 A	09-09-1998
		JP 2000500632 T	18-01-2000
WO 9821688 A	22-05-1998	EP 0976089 A	02-02-2000

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☒ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☐ **FADED TEXT OR DRAWING**
- ☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☒ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.